JSM2007

July 29 – August 2, 2007

Abstract Book

Table of Contents

UTAH

Salt Lake City

July 29 – August 2, 2007

Session **Abstracts**

1 Bayesian Statistics and Random Matrices ✪

Section on Bayesian Statistical Science

Sunday, July 29, 2:00 pm–3:50 pm

Covariance Selection and Bayes Classification via Modal Shrinkage Estimators

❖Jingqin Luo, Washington University in St. Louis, Box 8067, Division of Biostat, School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110, *rosy@wubios.wustl.edu*; Merlise A. Clyde, Duke University; Edwin Iversen, Duke University

Key Words: Bayesian analysis, covariance selection, Baye classification, shrinkage regression, scale mixtures of normals

Due to the positive definiteness constraint and the rapidly growing number of parameters with dimensions, covariance estimation in a multivariate normal population has been a classic but challenging statistical problem. Many approaches shrink a covariance/precision matrix toward some special parsimonious structures, which may suffer from misspecification error. By describing the covariance selection problem as a system of linear recursive equations, we work in the Cholesky decomposition framework of a precision matrix. Through application of Bayesian shrinkage regressions, we obtain robust estimators for a precision matrix of a flexible sparse pattern. A further application of Bayesian shrinkage regressions to Bayes classifier results in classifications comparable to some state-of-the-art methods.

Discrete Hierarchical Model Search with Conjugate Priors for Log-Linear Parameters

❖Adrian Dobra, University of Washington, 15514 Bothell Way NE, Apt. 202, Seattle, WA 98155, *adobra@u.washington.edu*; Helene M. Massam, York University

Key Words: Model selection, log-linear models, Bayesian statistics, Contingency tables, Reversible jump

We propose a novel reversible jump Markov chain Monte Carlo method (Green, 1995) for model selection in hierarchical log-linear models. Our approach is based on the conjugate priors for log-linear parameters introduced in Liu and Massam, 2006. We discuss the computation of Bayes factors as an alternative to RJMCMC and the Bayesian iterative proportional fitting algorithm for sampling model parameters. We also compare our work with similar results based on multivariate normal priors for log-linear models (Dellaportas and Forster, 1999; King and Brooks, 2001).

Local Graphical Model Search

❖Liang Zhang, Duke University, 214 Old Chem, ISDS, Durham, NC 27708, *lz9@stat.duke.edu*; Adrian Dobra, University of Washington; Mike West, Duke University

Key Words: Local Graphical Model Search, MCMC, Metropolis-Hastings

Compared to general graphical model search methods, we present here our work on local graphical model search algorithms. Local graphical model search algorithms will apply to the problem if we are only interested in one gene Y in thousands of genes in the gene expression data, for example, and wish to understand the graphical structure of Y and its graphical structure, where usual (global) graphical model search methods will not be efficient and precise. Also, the prediction of Y based on the local graphical structure is one of our interests. Monte Carlo Markov Chain methods and Shotgun Stochastic Search will be tried to do the search. We will provide several examples to analyze the efficiency as well as the precision of our algorithms.

Flexible Priors for Covariance Matrices in Gaussian Graphical Models

❖Bala Rajaratnam, Statistical and Applied Mathematical Sciences Institute, 19 T.W. Alexander Dr., Durham, NC 27709, *br62@samsi.info*; Carlos Carvalho, Duke University; Helene M. Massam, York University

The hyper-inverse Wishart distribution is a commonly used prior for Bayesian inference on covariance matrices in Gaussian Graphical models. This prior has the distinct advantage that it is the conjugate prior for this model but suffers from lack of flexibility in high dimensional problems due to its single shape parameter. In this paper, we propose flexible classes of priors for posterior inference on covariance matrices in Gaussian Graphical models that allow for up to k+1 shape parameters where k denotes the number of cliques in the graph. We investigate the corresponding Bayes estimators under usual losses considered in the literature and exploit the conjugacy relationship in these models to express Bayes estimators in closed form. The closed form solutions allow us to avoid heavy computational costs that are usually incurred in these problems.

2 The Special Interest Group on Volunteerism: Two Years Out

Section on Statistical Education, Committee on Scientific Freedom and Human Rights, Section on Government Statistics

Sunday, July 29, 2:00 pm–3:50 pm

Update on SIG on Volunteering Efforts

❖Robert Starbuck, Wyeth, 500 Arcola Road, A-5114A, Collegeville, PA 19426-3930, *starbur@wyeth.com*

A number of volunteering efforts are underway and being tracked and managed by the Special Interest Group on Volunteering. An update on the progress of these efforts will be presented, including a brief description of the efforts, and a mention of the person leading the effort and the key participants.

Developing Graduate Degrees in Statistics at Universities in Vietnam

❖Alicia Carriquiry, Iowa State University, Department of Statistics, 102D Snedecor Hall, Ames, IA 50011-1210, *alicia@iastate.edu*

Key Words: MS degree, Vietnam

This work is the joint effort of the Committee for Graduate Education in Vietnam, established under the umbrella of the SIG in Volunteerism of the ASA. The committee includes 14 members from academia and industry and was charged with developing approaches to create graduate degree programs in statistics at universities in Vietnam. The committee held several electronic discussions and prepared a report proposing a plan for a MS degree in statistics. The components of the MS degree in statistics, the resources that must be in place to establish the new degree program and possible strategies for building the necessary capacity are discussed in the report. The committee is currently constructing a web-based resource with course notes, datasets and software. Findings and web-based resources can be adapted for implementation in other developing nations.

Discussion of the Special Interest Group on Statistical Volunteerism

❖Jana Asher, Carnegie Mellon University, 26 Underwood Place NE, Washington, DC 20012, *jana@asher-resnick.us*

Key Words: Volunteerism, Human Rights, Humanitarian

The Special Interest Group on Statistical Volunteerism is now two years old and going strong. A variety of initiatives - from a meta-analysis of data from Darfur to the development of a protocol for rapid response to requests for statistical assistance in humanitarian and human rights crises - have been initiated. This paper discusses the initiatives of the Special Interest Group from the viewpoint of an active participant in the international human rights regime and makes suggestions for future directions for the Group.

Promoting Quantitative Literacy for School-Age Children

❖Shail J. Butani, Bureau of Labor Statistics, 6007 Marilyn Dr., Alexandria, VA 22310, *butani.shail@bls.gov*

Key Words: Quantitaive Literacy

These are some of the activities I have undertaken for Washington Statistical Society to promote quantitative literacy (QL). A vision of the Special Interest Group of volunteers is to inspire other statisticians to adopt these or similar ones in their local communities. 1) Conducted and organized speakers and materials for career days for over 100 math classes each year. 2) Participated and provided consultants for QL workshops conducted by ASA for local teachers. 3) Provided statisticians to assist in developing math curricula for Fairfax County Public Schools. 4) Conducted and provided statisticians for elementary schools teachers' workshops. 5) Presented materials at Female Achieving Mathematics Equity (FAME) project. 6) Provided speakers for Girls Excelling in Math and Science (GEMS) programs. 7) Conducted and provided consultants for girl scouts' workshops.

3 Fundamental Statistical Questions Underlying Clinical Trials ● ✪

Biopharmaceutical Section, ENAR, Biometrics Section, WNAR

Sunday, July 29, 2:00 pm–3:50 pm

Beyond Statistical Methods: Quality of Trial Issues

❖Guoxing Soon, Food and Drug Administration, MD , *soong@cder.fda.gov*

Key Words: missing data, open label, quality of trial

Statisticians are sometimes being referred to as number crunchers. This is unfortunate and in no way reflect the real role of statisticians in clinical research. Although statistical methods are important, what gets into those formula are even more important and should be an integral part of statistical considerations. For example, a lot of papers have been written regarding the missing data, however, no statistical method can truly address this issue satisfactorily because we can not make unknowns known afterward. A better approach is to think creatively how to reduce the missing rate and how to collect information that will help addressing this issue. In this talk I will go through some trial design, conduct and data collecting issues and discuss how statistician can play an important role in these areas.

On Criteria for a Measure of Statistical Evidence in Clinical Trials: What We Want and What We Don't

❖Jeffrey D. Blume, Brown University, Center for Statistical Sciences, Box G-S121, Providence, RI 02818, *jblume@stat.brown.edu*

Key Words: Statistical Evidence, Clinical Trials

It is important to assess the strength of statistical evidence in a clinical trial. But how statisticians choose to communicate this varies by philosophy and training. Some argue that this is communicated by the p-value, some argue for the posterior probability and some for Likelihood ratios. We propose three evidential quantities that every clinical trial should present: (1) the probability that the study design will generate misleading evidence, (2) the strength of the evidence in the observed data, and (3) the probability that the observed evidence is misleading. We'll argue that no matter what philosophical approach one prefers, only #2 and #3 are relevant at the end of the study and that #1 is always critical in the design stage. We'll explore if current paradigms identify these quantities and why this framework resolves problems of multiple comparisons and multiple looks.

4 Some New Developments in Statistical Learning ● ✪

Section on Statistical Computing, Section on Bayesian Statistical Science, Section on Teaching Statistics in the Health Sciences, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences

Sunday, July 29, 2:00 pm–3:50 pm

The Adaptive Lasso and Its Oracle Properties

❖Hui Zou, The University of Minnesota, 362 Ford Hall, 224 Church Street, Minneapolis, MN 55455, *hzou@stat.umn.edu*

Key Words: Adaptive Lasso, Oracle Propertties, Model Selection

The lasso is famous for its ability to automatically produce a sparse subset model. The LARS algorithm further facilitates the applications of the lasso in practice. The lasso method now seems to become the default choice for building a sparse model. In this talk, we first show a necessary condition for the lasso variable selection to be consistent. We present some examples in which the lasso is inconsistent for variable selection. We then propose a new version of the lasso, called the adaptive lasso, where adaptive weights are used for penalizing different coefficients in the lasso penalty. We show that the adaptive lasso is consistent in variable selection, and in addition, it possesses the oracle properties.

Robust Support Vector Machines

❖Yufeng Liu, The University of North Carolina at Chapel Hill, 306 Smith Building, CB 3260, Chapel Hill, NC 27599, *yfliu@email.unc.edu*

Key Words: classification, D.C. Algorithm, Fisher Consistency, Regularization, Truncation

The Support Vector Machine (SVM) has been widely applied for classification problems in both machine learning and statistics. Despite its popularity, it still has some drawbacks in certain situations. In particular, the SVM classifier may be sensitive to outliers in the training sample. Moreover, the number of support vectors (SVs) can be large in many applications. To circumvent these drawbacks, we propose the robust truncated-hinge-loss SVM (RSVM), which utilizes a truncated hinge loss. The RSVM is shown to be more robust to outliers and deliver more accurate classifiers using a smaller set of SVs than the standard SVM. Our theoretical results show that the RSVM is Fisher consistent, even when there is no dominating class, a scenario that is particularly challenging for multicategory classification.

Bayesian Ensemble Active Learning

❖Hugh Chipman, Acadia University, Department of Mathematics and Statistics, Wolfville, NS B4P2R6 Canada, *hugh.chipman@gmail.com*; Edward I. George, University of Pennsylvania; Robert McCulloch, The University of Chicago Graduate School of Business

Key Words: sequential design, machine learning, flexible modelling, Bayesian, active learning, ensemble

In supervised learning problems, ``Active Learning” refers to the iterative process of sequential data selection and model building, with the goal of building a better model while requiring fewer observations. From a statistical viewpoint, sequential design of experiments seeks to solve a similar problem, but often focuses on parametric models such as linear regression. We consider the challenge of Active Learning with very flexible models. In order to meet this challenge, a framework for formal statistical inference must be available for the model. We consider Bayesian Additive Regression Trees, a flexible ensemble model which can deal with high dimensionality, irrelevant predictors, nonlinear relationships, interactions, and local effects. We will discuss issues involved in the development of BART as an active learning tool, including computation and comparisons with other methods.

5

Time Series Studies of Air Pollution and Health ●

WNAR, ENAR, Section on Health Policy Statistics, Section on Risk Analysis, Biometrics Section, Section on Statistics and the Environment

Sunday, July 29, 2:00 pm–3:50 pm

Estimating the Distributed Lag between Air Pollution and Hospitalization Using a Bayesian Hierarchical Model

❖Roger Peng, Johns Hopkins University, Department of Biostatistics, 615 N Wolfe St, Baltimore, MD 21205, *rpeng@jhsph.edu*; Francesca Dominici, Johns Hopkins University; Leah Welty, Northwestern University

Key Words: air pollution, time series, distributed lag, hierarchical model

Numerous epidemiologic investigations have provided strong evidence of an association between increased levels of ambient air pollution and increased levels of hospital admissions. Time series studies have demonstrated an association typically at 0, 1, or 2 days after an air pollution episode. An important research aim is to obtain a more detailed understanding of the time course between exposure to air pollution and the onset of disease. Information about this time course could provide the basis for hypotheses concerning the mechanism by which air pollution causes disease. We introduce an hierarchical Bayesian distributed lag model that incorporates prior information about the time course of pollution effects and information from multiple locations. We examine the effects of pollution on cardiorespiratory diseases in a large United States air pollution and hospitalization database.

Reassessing the Relationship Between Ozone and Short-Term Mortality in U.S. Urban Communities

❖Richard L. Smith, The University of North Carolina at Chapel Hill, Department of Statistics and Operations Research, Chapel Hill, NC 27599-3260, *rls@email.unc.edu*; Baowei Xu, The University of North Carolina at Chapel Hill; Allen S. Lefohn, A.S.L. & Associates; Paul Switzer, Stanford University

Key Words: Air pollution and mortality, Hierarchical models, Nonlinear regression, Time series analysis

In this paper we re-analyze data collected in connection with the National Morbidity, Mortality and Air Pollution Study that have led to estimates of the relationship between ozone and short-term mortality, with the intention of examining the sensitivity of the results with respect to a number of methodological features. Among the features we examine are (a) sensitivity of the results to the selection of lags for meteorological and air pollution variables, including backwards in time analysis; (b) nonlinear models for the ozone-mortality relationship; (c) alternative approaches to distributed lags; (d) alternative approaches to hierarchical models for combining data from different cities.

Looking Through the Time Series Study Lens: Do We See Reality or Fiction?

❖Elizabeth A. Sheppard, University of Washington, Box 357232, Seattle, WA 98195-7232, *sheppard@u.washington.edu*

Key Words: air pollution, times series study, ecological study, study design

Air pollution scientists and policymakers rely heavily on time series study estimates of air pollution health effects. As a design time series studies are a paradox. By comparing day-to-day changes in air pollution concentrations with health event counts, their ecological structure suggests loss of information and bias in the parameter estimates. However, because they rely on huge populations from entire metropolitan areas, such studies are able to detect very small effects. In this talk I discuss the utility and interpretation of time series studies. I frame the discussion in terms of an individual-level disease model to describe the association between personal air pollution exposure and health. I discuss the impact of the ecological design, the specific risk and exposure models, and the need to substitute ambient concentration for personal exposure in the time series analysis.

6

Threshold Estimation for Functions in Nonparametric and Semiparametric Pproblems ✪

Section on Nonparametric Statistics

Sunday, July 29, 2:00 pm–3:50 pm

Asymptotically Optimal Tests Under Loss of Identifiability in Semiparametric Models

Rui Song, The University of North Carolina at Chapel Hill; ❖Michael Kosorok, The University of North Carolina at Chapel Hill, Department of Biostatistics, Chapel Hill, NC 27599-7420, *kosorok@unc.edu*; Jason Fine, University of Wisconsin-Madison

Key Words: Semiparametric methods, Optimal tests, Loss of identifiability, Change-point models, Transformation models, Mixture models

We consider tests of hypotheses when the parameters are not identifiable under the null in semiparametric models. Under a weighted average power criterion, exponential average tests are characterized and shown to be asymptotically optimal. The results can be applied to a variety of semiparametric models, for example, tests of presence of change-point in transformation models, tests of regression parameters in gamma frailty models, and tests of the number of mixture components in finite components mixture models are discussed. We also propose a modified weighted bootstrap for computing the critical values of the test statistic.

Trajectories as Predictors of Univariate Responses

❖Ian McKeague, Columbia University, , *im2131@columbia.edu*

Key Words: functional data analysis, change-points, empirical processes, misspecified models, diffusion processes, functional linear regression

This talk discusses a new type of functional regression model involving trajectories as predictors of univariate responses. The basic idea is to include time points among the parameters of interest, and to fit a (possibly mis-specified) model using just the values of the trajectories at those times.

This provides greater interpretability than the functional linear regression approach that has become popular in recent years. The new methodology involves nonstandard asymptotic theory; specifically, non-Gaussian limit distributions and rates of convergence that are determined in some sense by the smoothness of the trajectories. In the case of trajectories that are single-jump counting processes, it specializes to an existing type of change-point analysis. Trajectories of diffusion-type are also studied in detail.

Multistage Procedures for Change Point Estimation

❖George Michailidis, The University of Michigan, 439 West Hall, 1085 S. University, Ann Arbor, MI 48109, *gmichail@umich.edu*

Consider a constant regression model for a bounded covariate that has a single discontinuity (change point). It is assumed that one can sample the covariate at different values and measure the corresponding responses. Budget constraints dictate that a total of n such measurements can be obtained. The goal is to estimate accurately the location of the change-point. A multistage procedure is proposed and its properties examined, where at each stage an estimate of the change point is obtained and new points are sampled from its neighborhood. The asymptotic distribution of the least squares estimate is derived using ideas from empirical processes. The improved efficiency of the procedure is demonstrated using real and synthetic data. The problem is primarily motivated by problems in engineering systems.

7 Medallion Lecture I

IMS, General Methodology, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

Modeling Genes: Statistical Challenges in Modern Genetics

❖Peter Donnelly, University of Oxford, Department of Statistics, 1 South Parks Road, Oxford, International OX1 3TG United Kingdom, *donnelly@stats.ox.ac.uk*

Key Words: statistical genetics, association studies, coalescent, human diseases

Advances in experimental technology have made possible hypothesis-free, genome-wide, case-control association studies. These are proving to be a very powerful tool for unlocking the genetic basis of common human diseases, but they bring with them a number of challenging statistical problems and opportunities. The talk will describe some of these in the context of the largest such study to date, the Wellcome Trust Case Control Consortium, a UK collaboration studying the genetic basis of seven common human diseases which measured the genetic types of 2000 cases for each disease and 3000 common controls at each of 500,000 positions in the human genome. The study identified more than 10 novel genes involved in the etiology of the diseases studied.

8 Nonlinear and Nonstationary Time Series Analysis ✪

International Indian Statistical Association

Sunday, July 29, 2:00 pm–3:50 pm

Multiple Local Whittle Estimation in Stationary Systems

❖Peter M. Robinson, London School of Economics, Houghton Street, London, WC2A 2AE United Kingdom, *p.m.robinson@lse.ac.uk*

Key Words: long memory, semiparametric estimation, cointegration, phase, consistency, asymptotic normality

Related long memory time series x(t), y(t) may have different or identical memory parameters. In the latter case the possibility exists of domination by a single common component, such that an instantaneous linear combination y(t)-flx(t) of the series has shorter memory than x(t), y(t). A semiparametric version of the problem presents four unknown parameters of interest: two memory parameters, a phase parameter, and fl. We estimate them jointly by optimizing a local Whittle function that entails a single smoothing number, covering also the case where fl=0 and y(t) and z(t) have different memory parameters. The consistency proof is non-standard in that the fl estimate converges faster than the others. Joint asymptotic normality is also established. Misspecification of the phase can slow convergence of estimates of fl, and cause inconsistency of memory parameter estimates.

Generalized R-estimators Under Conditional Heteroscedasticity

❖Kanchan Mukherjee, The University of Liverpool, Department of Mathematical Sciences, Peach Street, Liverpool, L69 7ZL United Kingdom, *k.mukherjee@liverpool.ac.uk*

Key Words: R-estimation, Autoregression, ARCH model, Heteroscedasticity

In this talk, we discuss a class of rank estimators of the parameters associated with the conditional mean function of an autoregressive model with heteroscedastic errors. The estimators are defined through a three-steps procedure. We discuss the role of these estimators in mis-specified models and their asymptotic distributions. The class of models includes Engel's ARCH model and the threshold heteroscedastic model. The class of estimators includes an extension of Wilcoxon-type rank estimator. Some simulation studies are presented.

Conditions for the Propagation of Memory Parameter from Durations to Counts and Realized Volatility

❖Clifford M. Hurvich, New York University, IOMS Dept. 44 West 4'th Street, Stern School of Business, New York, NY 10012, *churvich@stern.nyu.edu*

Key Words: Long Memory Stochastic Duration, Autoregressive Conditional Duration

We establish sufficient conditions on durations that are stationary with finite variance and memory parameter d in [0,1/2) to ensure that the corresponding counting process N(t) satisfies Var N(t) ~ C t**(2d+1) as t-> infinity, (C>0), so that the memory parameter in durations propagates to the same memory parameter in counts. We then show that any Autoregressive Conditional Duration ACD(1,1) model with a sufficient number of finite moments yields short memory in counts, while any Long Memory Stochastic Duration model with d>0 and all finite moments yields long memory in counts, with the same d. Next, we motivate the need for long memory: aggregation ultimately destroys all autocorrelation in counts, if and only if the counts have short memory. Finally, for a simple pure-jump price process, we show that the memory parameter in durations propagates all the way to the realized volatility.

Break Detection for a Class of Nonlinear Time Series Models

❖Richard A. Davis, Colorado State University, Department of Statistics, Ft. Collins, CO 80523, *rdavis@stat.colostate.edu*; Thomas C.M. Lee, The Chinese University of Hong Kong; Gabriel Rodriguez-Yam, Universidad Autonoma Chapingo

Key Words: nonlinear time series, structural break, MDL, genetic algorithm

We consider the problem of detecting break points for a broad class of nonstationary time series models. In this formulation, the number and locations of the break points are assumed to be unknown. Each segment of the data is assumed to follow a model from a family of nonlinear time series models. The minimum description length is used as a criterion for estimating the number of break points, the location of break points, and the parameters of the model in each segment. The best segmentation, found by minimizing MDL, is obtained using a genetic algorithm. The implementation of this approach is illustrated using a GARCH model, a stochastic volatility model, and a generalized state-space model as the base parametric model for the segments. Empirical results show the good performance of the estimates of the number of breaks and their locations for these various models.

9

Propensity-Score Matching for Observational Studies: Recent Extensions and Refinements ●

IMS, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Biometrics Section, WNAR

Sunday, July 29, 2:00 pm–3:50 pm

Semiparametric Sequential Estimators for Dynamic Causal Models

❖Michael H. Lechner, University of St. Gallen, Bodanstr. 8, SIAW, St. Gallen, 9000 Switzerland, *Michael.Lechner@unisg.ch*

Key Words: Dynamic treatment effects, causal effects, nonparametric identification, sequential randomisation, programme evaluation, panel data

This paper proposes inverse selection probability weighting and sequential propensity score matching to estimate dynamic causal effects. A Monte Carlo study shows that the suggested matching estimators perform well in small and medium sized samples. A re-evaluation of some parts of the Swiss active labor market policies based on a dynamic causal model allows the discussion of issues concerning the implementation of the new methods.

Propensity Score–Matching Design in Studies with Repeated Measures

❖Bo Lu, The Ohio State University, B110 Starling-Loving Hall, SPH, 320 West 10th Ave., Columbus, OH 43235, *blu@sph.osu.edu*

Key Words: Observatioanl Studies, Repeated Measures, Propensity Scores, Optimal Matching, Non-bipartite Matching

In observational studies with repeated measures, treatment effects need to be clearly defined. To obtain unbiased the treatment effect estimation, with potentially time-dependent treatment and covariates, it is desirable to balance the distribution of the covariates at every time point. A propensity score based on the Cox proportional hazards model is proposed to incorporate the time dependent feature and implemented in risk set matching. Different optimal matching designs have been explored under different study setups. Simulation results and analysis of a practical example are presented.

Matching with Propensity and Prognostic Scores

❖Ben B. Hansen, University of Michigan, Statistics Department, 439 West Hall, Ann Arbor, MI 48104, *ben.hansen@umich.edu*

Key Words: covariate balance, multivariate distance, subclassification, poststratification

Matching is among the oldest and most transparent methods of adjustment for observational studies. To extract treatment effects, it says, first match treated subjects to untreated but otherwise similar controls; then compare subjects to their matched counterparts and aggregate these individualized comparisons. Used with propensity scores, the method is simple and far-reaching, addressing high-dimensional covariates and helpfully separating outcome analysis from the process by which candidate matches are developed and evaluated. These benefits are shared with prognostic scores, cousins of propensity scores addressing important aspects of the data that propensities ignore. Combining propensity and prognostic scores enhances matching's scope and power as compared to matching on propensities only, while avoiding hazards to which earlier prognostic-matching proposals have succumbed.

10

Inference for Multistate Data Under Complex Censoring Structures ●

ENAR, Biometrics Section, Section on Physical and Engineering Sciences

Sunday, July 29, 2:00 pm–3:50 pm

Statistical Analysis of Multivariate Interval-Censored Data

❖Jianguo Sun, University of Missouri-Columbia, Department of Statistics, Columbia, MO 65211, *sunj@missouri.edu*; Manhua Chen, University of Missouri-Columbia; Xingwei Tong, Academia of Science

Key Words: AIDS studies, Correlated failure times, medical studies, regression analysis

Multivariate interval-censored failure time data occur when there exist several correlated failure times of interest and for each failure time, only interval-censored observations are available. Such data can arise from many fields including medical follow-up studies, psychological studies and tumorgenicity experiments. This talk will discuss statistical analysis of these data and in particular, their regression analysis under several commonly used regression models will be investigated.

Analysis of Longitudinal and Response/Failure Time Data from an Observational Study with Missing Observations

❖Dianne M. Finkelstein, Harvard/MGH, 50 Staniford Street, Suite 560, Boston, MA 02114, *dfinkelstein@partners.org*; Natasa Rajicic, Pfizer Inc.; David A. Schoenfeld, MGH/Harvard

Key Words: interval censored, failure time, multi-state

Observational studies often are interested in the relationship between a longitudinal marker and clinical improvement (e.g., response) or progression. Often both the marker and the determination of clinical status are assessed at regular clinic visits. If a subject misses a visit, the response and/or progression time will only be known to have occurred in the interval bounded by the last negative, and first positive visit times (interval censored event time data). In addition, the marker data will be missing. This talk will focus on methods for analysis of data from such a study. The methods are discussed in the context of data from a diabetes and a trauma study.

Multiply Robust Higher-Order U-Statistics Estimators for Continuous-Time Multistate Data Structures

❖James Robins, Harvard School of Public Health, 677 Huntington Ave, Suit 821, Boston, MA 02115, *robins@hsph.harvard.edu*

Key Words: Nonparmaetric, Semiparametric, Dependent Censoring, U-statistics

Under coarsening at random, doubly robust estimators exist for the parameters of semiparametric models for continuous time multistate longitudinal data processes in the presence of dependent censoring explainable by high dimensional time-dependent covariate processes. In this talk, we describe new multiply (i.e., triply, quadruply, etc.) robust estimators based on higher order influence functions. These estimators are higher order U statistics. We posit a number of models for the joint distribution of the full data and for the censoring mechanism such that if any one of the many models is correct, we obtain a consistent asymptotically normal estimator of the parameter of interest. The more models one uses the higher the required order of the U-statistic.

11 Alzheimer's Disease Research in Utah: Analysis Methods for Early Detection, Epidemiology, and Clinical Trials ● ✪

ASA, Utah Chapter, Council of Chapters, ENAR, Section on Statistics in Epidemiology, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

Cognitive Change as a Biomarker in the Early Detection of Alzheimer's Disease

❖ Gordon J. Chelune, The University of Utah, Center for Alzheimer's Care, Imaging and Research, 650 Komas Drive, Ste 106A, Salt Lake City, UT 84108, *gordon.chelune@hsc.utah.edu*

Key Words: Reliable Change Methods, Individual Trajectories, Dementia, Cognitive Assessment of Change, norms for change

Alzheimer's Disease (AD) is characterized by progressive cognitive impairment. For new drug therapies to be maximally effective, early detection is critical. Single-point assessments using cohort norms are often insensitive early in the disease course as subtle neurocognitive changes may be masked by the effects of normal aging. Increasingly, there is reliance on serial testing to identify abnormal cognitive trajectories. However, norms for cognitive change do not exist, and there is a paucity of objective methods to characterize patterns of reliable change at the level of the individual. Regression-based reliable change methods may provide a means for identifying abnormal cognitive trajectories across serial assessments early in the disease course, even when the individual is still considered to be within the normal range, thus providing a potentially useful biomarker of prodromal AD.

Analyzing Repeated Neuropsychiatric, Cognitive, and Functional Measures in a Population-Based Study of Dementia

❖ Chris Corcoran, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, *corcoran@math.usu.edu*; Carl F. Pieper, Duke University; Peter Zandi, Johns Hopkins University

Key Words: repeated measures, dementia, cognition, multivariate models, latent class

Studies of elderly dementia patients incorporate a variety of measures to characterize the course of dementia. These metrics can be divided into three general categories - cognitive, behavioral, and functional - each consisting of subscales designed to capture more specific aspects of individual symptoms. Understanding what influences rates of change across these domains, and how these rates interact over time, may assist in creating and improving treatments. We illustrate the issues relative to analyzing repeated-measures data across multiple domains with data from the Cache County Memory Study - one of the only existing population-based longitudinal studies of dementia. In particular, we discuss and compare the use of multivariate hierarchical linear models and latent class trajectory analyses to assess patterns of change and correlations between trajectories across domains.

Responder versus Means Analysis: Illustrated with Data from a Phase 2 Study of Tarenflurbil (MPC-7869) in Alzheimer's Disease (AD)

❖ Kenton Zavitz, Myriad Pharmaceuticals, 320 Wakara Way, Salt Lake City, UT 84108, *kzavitz@myriad.com*; Mark Laughlin, Myriad Pharmaceuticals; Edward Swabb, Myriad Pharmaceuticals

Key Words: Responder Analysis, Power, Clinical Trials, Dementia, Simulation Study

Preclinical data suggests a potential for MPC-7869—a Selective Afl42-Lowering Agent (SALA)—to have disease-modifying properties. In subjects with mild AD on 800 mg BID, statistical significance was observed for a means analysis at 12 and 24 months on the ADCS-ADL and the CDR-sb with positive trends on the ADAS-cog. Post-hoc responder analyses showed similar results. Responder analyses are often preferred by clinicians because they reflect a patient oriented approach, however, this analysis loses power by categorizing a continuum. Statisticians often prefer the potentially more sensitive means comparisons. In order to explore the statistical relationship between these two analysis approaches, data was simulated using four different distributions (a normal and a bimodal distribution each with 2 standard deviations) and the power and efficiency of these two approaches were compared.

12 Measuring and Improving Data Quality: A Prerequisite to Harnessing the Power of Information ● ✪

Social Statistics Section, Section on Government Statistics, Section on Survey Research Methods

Sunday, July 29, 2:00 pm–3:50 pm

Assessing the Effectiveness of Using Panel Respondent Mode Preferences

❖ Thomas B. Hoffer, NORC at the University of Chicago, 1155 E. 60th Street, Chicago, IL 60637, *hoffer-tom@norc.uchicago.edu*; Karen Grigorian, NORC at the University of Chicago; Ron Fecso, National Science Foundation

Key Words: respondent mode preferences, multimode surveys, response rates, data quality

This paper uses data from the Survey of Doctorate Recipients (SDR), an ongoing panel survey of doctoral scientists and engineers, to assess whether and how mode preference data should be used to improve panel survey participation and data quality. In the 2003 SDR, respondents were asked to indicate whether they would prefer to complete via paper, Web, or CATI in the next (2006) round. Most indicated a preference for the Web mode but honoring that preference and the supporting system of e-mail prompts runs a risk of less-cooperative subjects more easily deflecting prompts. Such cases may then wind up in the relatively expensive CATI mode, and defeat the purpose of honoring preferences. We analyze response rate and quality indicators from the 2001, 2003, and 2006 surveys to assess the extent to which honoring respondent mode preferences had a positive or negative effect on survey outcomes.

Look and Listen, But Don't Stop: Interviewers and Data Quality in the 2007 SCF

❖Arthur B. Kennickell, Federal Reserve Board, Mail Stop 153, Washington, DC 20551, *Arthur.Kennickell@frb.gov*

Key Words: Data quality, Interviewers, Quality control

In most field surveys, the data collection process is observed only by the respondents and the interviewers. Others can observe only the traces reflected in the data and paradata. Careful hiring and training of interviewers is, of course, very important in maintaining data quality. But creating a continuing mechanism for clarifying, reinforcing and enforcing the survey protocols is also important. The 2004 Survey of Consumer Finances (SCF) introduced a new two-part system of data review coupled with regular feedback to the interviewers throughout the field period. Based on the experience with that system, a more refined process was developed for the 2007 SCF. This paper presents a discussion of the process as seen from the perspective of the ongoing survey.

Survey Quality and Quality Metrics: An Economic Approach

❖Julia Lane, NORC at the University of Chicago, 55 E Monroe St, Suite 4800, Chicago, IL 60603, *gasparac-julie@norc.org*; Suzanne Bard, NORC at the University of Chicago; Judith Petty, NORC at the University of Chicago

Key Words: Quality metrics, Survey quality, Economic Approach

This paper describes the results of a new initiative at NORC that directly includes quality metrics into a survey instrument. The economic approach used is a principal/agent model, which argues that wages, together with alternative employment opportunities, affect the quality of agents that are available to a firm (the principal). The paper uses these direct quality metrics measures of field interviewer quality, and examines the impact of alternative employment opportunities on the recruitment and retention of high quality interviewers. The analysis is based on the Survey of Consumer Finances.

13 2007 Government Statistics, Social Statistics, and Survey Research Methods Student Paper Awards Session ●

Section on Survey Research Methods, Social Statistics Section, Section on Government Statistics, SSC

Sunday, July 29, 2:00 pm–3:50 pm

Small-Area Estimation in Multivariate Surveys

❖Hukum Chandra, University of Southampton, Building 39, Southampton, SO1 71BJ United Kingdom, *hchandra@soton.ac.uk*; Ray Chambers, University of Wollongong

Key Words: Multivariate Surveys, Multipurpose sample weights, MBD approach, Small Areas

Sample surveys are generally multivariate, in the sense that they measure more than one response variable. In theory, each variable can then be assigned an optimal weight for estimation purposes. However, it is often a distinct practical advantage to have a single weight that is used with all variables collected in the survey. This paper describes how such multipurpose sample weights can be constructed when small area estimates of the survey variables are required. The approach is based on the model-based direct (MBD) method of small area estimation described in Chambers and Chandra (2006). Empirical results show that MBD estimators for small areas based on multipurpose weights perform well across a range of variables that are often of interest in business surveys. Furthermore, these results show that the proposed approach is robust to model misspecification.

Nonparametric Modeling of the Second Order Structure of Processes with Time-Varying Memory

❖Georg Goerg, Vienna University of Technology, Argentinierstrabe 8/105, Vienna, 1040, Austria, *e0225792@stud3.tuwien.ac.at*; Dana Draghicescu, Hunter College

Key Words: Hurst parameter, nonstationary time series, time-varying memory, optimal bandwidth selection, returns, river flow

In this paper we consider nonstationary time series, characterized by time-varying memory. We model the temporal changes of the memory parameter (or Hurst parameter) in the time domain, by using a moving window approach. Our estimation procedure incorporates a data-driven scheme for optimal bandwidth selection. The proposed methodology is illustrated on hydrological and financial datasets.

Cosponsorship in the U.S. Senate: A Mixed Effects Approach to Detecting the Subtle Influence of Social Relational Factors on Legislative Behavior

❖Justin Gross, Carnegie Mellon University, 984 Greenfield Ave, Pittsburgh, PA 15217, *jhgross@andrew.cmu.edu*; Stephen Fienberg, Carnegie Mellon University; Cosma Shalizi, Carnegie Mellon University; David Krackhardt, Carnegie Mellon University

Key Words: Relational data, random effects, social network analysis, political methodology, GLMM, legislative behavior

Why do members of the United States Congress choose to cosponsor legislation proposed by their colleagues and what can we learn from their patterns of cosponsorship? These questions have attracted increased interest among legislative scholars over the past several years, and are, fundamentally, questions about relationships. Unfortunately, most methods of statistical inference with which political scientists are likely to be familiar tend to be ill-suited for the analysis of relational data, where observations are interdependent. We use mixed effects to capture network-type patterns of dependence (e.g., sender- and receiver-specific effects, reciprocity). Explicitly modeling network phenomena allows us to more confidently examine whether various dyadic properties of interest, such as similarities and opportunities for interaction between senators, translate into increased cosponsorship.

Imputing Missing Data in the FBI's Uniform Crime Reports

❖Clint Roberts, The Ohio State University, 2513 Mt. Holyoke Rd., Columbus, OH 43221, *roberts.694@osu.edu*; Elizabeth A. Stasny, The Ohio State University; Michael D. Maltz, The Ohio State University

Key Words: GLM, Poisson regression, Time series

The Uniform Crime Reports, collected by the FBI, contain monthly crime counts for each of the seven index crimes, but for one reason or another, a police agency may miss reporting for a particular month. The data are not complete, hence the need for the development of an imputation procedure to fill in the gaps. Since the early 1960's, an imputation technique implemented by the FBI has been used to make annual crime count estimates. This approach ignores concerns of seasonality and does not make use of the agencies' long-term data trends. Computing power has radically improved since the 1960's, and it is now feasible to develop a more involved yet flexible method that can incorporate more information into our estimation procedure. Our model-based approach also has the added value of making available variance estimates for the imputed data.

Using the Multiple Imputation Technique To Correct for Measurement Error and Statistical Disclosure Control in Sensitive Count Data in a National Survey

❖Mandi Yu, University of Michigan, 426 Thompson Street, Rm 4050, Program in Survey Methodology, Ann Arbor, MI 48106, *mandiyu@isr.umich.edu*

Key Words: Multiple Imputations, Sensitive Question, Measurement Error, Social Desirability, Statistical Disclosure Control, Bayesian

Measurement error in sensitive question is pervasive, therefore, biasing the estimation of most statistical models. The objective of this paper is to correct for measurement error in the number of life-time sexual partners by treating it as a missing data problem and using multiple imputation technique to synthesize this underlying "true" attribute. Bayesian Bivariate Poisson model with diffuse Gaussian priors was fitted to the 1996 General Social Survey combining knowledge of data quality from the mode experiment conducted by Tourangeau and Smith (1996). Ignored in existing literature, the threat of augmented disclosure harm from releasing both imputed and original data to the public was recognized and tackled by statistical perturbation. Bias reduction and statistical integrity were evaluated. Markov Chain Monte Carlo algorithm was programmed using WinBUGS.

14 Efficient Designs in Drug Development ● ✪

Biopharmaceutical Section, Biometrics Section, ENAR

Sunday, July 29, 2:00 pm–3:50 pm

Bayesian Forecasting of Multicenter Enrollment

❖Byron Jones, Pfizer Inc., Pfizer IPC 193, Ramsgate Road, Sandwich, Kent, International CT13 9NJ United Kingdom, *Byron.Jones@pfizer.com*

Key Words: multicenter, enrollment, Bayesian, prediction, updating, decision

A key design element of a multicenter trial is the number of centers and the rate at which each center enrolls patients. During the trial these rates are monitored to determine if the trial will complete enrollment on time. If enrollment is going too slowly then either more centers need be added or the rates of the existing centers increased. We describe (i) a Bayesian approach to monitoring enrollment and forecasting future rates and (ii) a decision scheme for deciding whether to add or drop centers or increase current enrollment rates.

A Bayesian Design and Analysis for Dose-Response Using Informative Prior Information

❖Michael Smith, Pfizer Inc., Ramsgate Rd, IPC096, Sandwich, CT13 9NJ United Kingdom, *mike.k.smith@pfizer.com*; Scott Marshall, Pfizer Inc.

Key Words: Dose-response, Bayesian, Emax model, Informative priors, relative potency, proof of concept

We wish to use prior information on an existing drug in the design and analysis of a dose-response study for a new drug candidate within the same pharmacological class. Using Bayesian methodology this prior information can be used quantitatively and the randomization can be weighted in favor of the new compound, where there is less information. An Emax model is used to describe the dose-response of the existing drug. The estimates from this model are used to provide informative prior information used for the design and analysis of the new study to establish the relative potency between the new compound and the existing drug therapy. Simulations were performed to assess the operating characteristics of various designs and robustness to underlying assumptions.

Model-Based Trial Design in Depression Using Bayesian Monitoring and Bootstrap Trial Simulation

❖Roberto Gomeni, GlaxoSmithKline, Via Fleming 4, Verona, 37135 Italy, *roberto.a.gomeni@gsk.com*

Key Words: Bayesian Monitoring, Posterior Probability of Superiority, Bootstrap, Trial Simulation, Model-based drug development, Risk analysis

We present a methodology for continuous monitoring of efficacy data (HAMD-17) in placebo-controlled trials in depression and to implement decision rules based on a Bayesian trial simulation to support progression or discontinuation of a treatment arm. The Posterior Probability of Superiority resulting from the application of MCMC to a longitudinal model of placebo and antidepressant drugs was used to estimates criteria to discontinue a treatment arm for futility, and to predict the treatment effect at study-end while the trial was still ongoing. The results showed the possibility to stop a trial for futility when about 50% of total information was available and to detect signal of a treatment effect based on limited information (<40%). Bootstrap analysis was able to protect from improper decision in presence of data heterogeneity, providing a framework for a scenario-based risk analysis.

Optimizing the Manufacturing of Phase III Material During Phase II Trials

❖Nitin Patel, Cytel Inc., 675 Massachusetts Ave, Cambridge, MA 02139, *info@cytel.com*; Francoise Vandenhende, Eli Lilly and Company

We will describe a Bayesian decision analysis model developed for late stage Phase II trials. The aim of the model is to explore the opportunity to reduce the time to market for a drug in a Phase II clinical trial by triggering the Compound Manufacturing and Control (CMC) process before completion of the trial. Since the CMC effort for a drug often lies on the critical path to release, early triggering has the promise of increasing revenues over the life of the drug. The risk is that the CMC work may be fruitless as there is a chance that the drug will fail on the Phase II trial. The model optimizes this trade-off and is currently being actively investigated for application to a Phase II trial of a cancer drug.

Accounting for Patient Heterogeneity in Phase II Clinical Trials

❖J. Kyle Wathen, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Box 447, Houston, TX 77030-4242, *jkwathen@mdanderson.org*; Peter F. Thall, The University of Texas M.D. Anderson Cancer Center; John D. Cook, The University of Texas M.D. Anderson Cancer Center

Key Words: Adaptive design, Bayesian design, Futility rule, Phase II clinical trial, Simulation

Phase II clinical trials typically are single-arm studies conducted to decide if an experimental treatment, E, is promising, relative to standard treatment, S, to warrant further investigation. Many methods exist for conducting phase II trials under the assumption of patient homogeneity. In the presence of patient heterogeneity, however, these designs are likely to draw incorrect conclusions. We propose a class of model-based Bayesian designs for single-arm phase II trials with a binary outcome and two or more prognostic subgroups. The designs' early stopping rules are subgroup-specific and allow the possibility of terminating some subgroups while continuing others. We require informative priors on S and subgroup main effects, and noninformative priors on E and treatment-subgroup interactions. An algorithm for computing prior hyperparameter values and a simulation study are presented.

15 Novel Analyses of Time-To-Event Data ● ✪

Biopharmaceutical Section, ENAR, Section on Risk Analysis
Sunday, July 29, 2:00 pm–3:50 pm

Extensions of Cure Models for Clustered Time-to-Event Data

❖Jeremy Taylor, University of Michigan, 1420 Washington Heights, Department of Biostatistics, Ann Arbor, MI 48109, *jmgt@umich.edu*; Yingwei Peng, Queen's University

Key Words: cure models, clustered data

Cure models are attractive approaches in survival analysis when there is a rationale for a non-susceptible group and there is good follow-up beyond the typical event times. The standard mixture cure model has a logistic model for long term incidence and a Cox model for latency amongst the susceptible group. We extend this model to the situation of clustered observations. In a conditional approach random effects are introduced in both the incidence model and the latency model to incorporate the cluster effects. In a marginal approach the variance-covariance of the estimates is obtained from a sandwich estimator. For both methods estimation is performed by extensions of the EM algorithm. Simulation studies are performed. The methods are applied to a multi-institutional study of local recurrence of tonsil cancer patients who received radiation therapy.

Constructing Approximate Confidence Bounds for the True Survival Curve Using the Kaplan-Meier Survival Function by Adding an Imaginary Failure or Eternal Survivor

❖Craig B. Borkowf, Centers for Disease Control and Prevention (CDC), 1600 Clifton Road NE, Mail Stop E48, Atlanta, GA 30333, *uzz3@cdc.gov*

Key Words: censored data, clinical trial, Clopper-Pearson method, Greenwood variance estimator, Kaplan-Meier survival function, survival analysis

We present a method for constructing confidence bounds for the true survival curve as estimated by the Kaplan-Meier survival function, derived as an extension of the Clopper-Pearson method for Binomial proportions. As a consequence of the definition of the Kaplan-Meier survival function, the upper bound should decrease only when events occur, but the lower bound should decrease when either events or censorings occur. The confidence interval width is proportional to the current weights of the observations still present in the sample under Efron's redistribute-to-the-right algorithm. Simulations show that the proposed method gives at least nominal coverage in a variety of situations, even for small sample sizes, except in the right tail of the survival distribution with heavy censoring. Finally, the proposed method is illustrated with an application from a leukemia clinical trial.

A Two-Stage Procedure for Comparing Hazard Rate Functions

❖Peihua Qiu, The University of Minnesota, 313 Ford Hall, 224 Church St. SE, Minneapolis, MN 55455, *qiu@stat.umn.edu*

Key Words: Additive tests, Censoring, Crossing hazard rates, Sequential tests, Significance level, Survival analysis

We consider comparison of two hazard rates that may cross each other. A number of existing procedures for handling this problem only consider the alternative hypothesis with crossing hazard rates; many other realistic cases are excluded from consideration. We propose a two-stage procedure that considers all possible alternatives, including ones with crossing or running parallel hazard rates. To define its significance level and p-value properly, a new procedure for handling the crossing hazard rates problem is suggested, which has the property that its test statistic is asymptotically independent of the test statistic of the logrank test. We show that the two-stage procedure, with the logrank test and the suggested procedure for handling the crossing hazard rates problem used in its two stages, perform well in applications in comparing two hazard rates.

Improved Hazard Ratio Estimation in Small Time-to-Event Trials

❖Devan V. Mehrotra, Merck & Co., Inc., Mailstop UG1CD44, 351 N Sumneytown Pike, North Wales, PA 19454, *devan_mehrotra@merck.com*; Arthur Roth, Pfizer Inc.

Key Words: Cox model, Efron approximation, Generalized logrank test, Proportional hazards, Survival Analysis, Ties

The Cox proportional hazards (CPH) model is the standard tool for hazard ratio estimation in time-to-event trials. Mehrotra and Roth (2001) proposed an alternate approach based on a generalized logrank (GLR) statistic, and showed it to be notably more efficient than CPH for small trials (< 200 patients), and equally efficient for large trials. However, a limitation of the GLR approach is that it assumes no (or very few) tied event times, which is restrictive because continuous event times are often rounded to the nearest day, week, etc., resulting in ties. The best tie-handling methods for the Cox model are those of Kalbfleisch and Prentice [KP] (1973) and Efron (1977). Two extensions of the original GLR approach, GLR-KP and GLR-E, will be developed and shown (via simulations) to yield estimators that are more efficient than those based on the KP and Efron methods, respectively.

16 Design and Inference in Multinational Trials ●

ENAR, Biometrics Section, Biopharmaceutical Section, WNAR
Sunday, July 29, 2:00 pm–3:50 pm

Design of Multinational Study for Global Development

❖Yoshiharu Horie, Nippon Boehringer Ingelheim Co. Ltd., Yato, Kawanishi, 666-0193 Japan, *horiey@kaw.boehringer-ingelheim.com*; Masahiro Takeuchi, Kitasato University

Key Words: Multinational study, Study design, Dose response, Drug development

One of the new challenging issues regarding global development in the ICH E5 guideline is to plan and conduct the multinational study and inference the results of study. A lot of papers for multinational studies were published in medical journals, especially cardiovascular disease. Overall results were attracted attention from investigators as the scale of that type of study was large. However, there are several considerations from a pharmaceutical point of view if results of each country are included into the new drug application in each country. We show these considerations and propose the corresponding approaches.

Sample Size Determination of Japanese Patients for Multiregional Clinical Trial in Oncology

❖ Risa Sekiguchi, Eli Lilly and Company, Sannomiya Plaza Bldg 7 1 5, Isogamidori Chuoku, Kobe, 651-0086 Japan, *sekiguchi_risa@lilly.com*; Sachio Ogawa, Eli Lilly and Company; Hiroyuki Uesaka, Eli Lilly and Company

Key Words: hazard ratio, multi-regional clinical trial, oncology, sample size

Recently, the pharmaceutical companies in Japan tend to join multiregional clinical trials in order to achieve a simultaneous submission globally. One of the critical issues in the design of a multi-regional trial is sample size determination of Japanese patients. It is not feasible to take account of statistical significance of efficacy in Japanese patients when we estimate the sample size of Japanese patients. It will be preferable to estimate the sample size of Japanese patients so that we can show that the effect size in Japanese patients is bigger than by the specified fraction of the effect size in the whole or in the patients other than Japan at least. We will present the example of how to estimate sample size of Japanese patients in multiregional clinical trial in oncology from this point of view.

Issues in Design of Multiregional Trials

❖ Yoichi Ii, Pfizer Japan Inc., Shinjuku Bunka Quint Bldg, 3 22 7 Yoyogi Shibuya ku, Tokyo, 151-8589 Japan, *yoichi.ii@pfizer.com*; Osamu Komiyama, Pfizer Japan Inc.; Kazuhiko Kuribayashi, Pfizer Japan Inc.

Key Words: Multi-regional trials, ICH-E5 Q&A 11, Multi-national trials, Bridging

ICH-E5 Q&A 11 introduced a concept of multi-regional trials (MRT) serving as a bridging study and intended to facilitate moving from “sequential bridging” to simultaneous global development. While it is still based on philosophy of ICH-E5, MRT is the first step toward “multi-national trials” that have been common in the US and EU, and try to talk multinational trials as a natural extension of ICH-E5. When planning a MRT, one needs to address design elements such as sample size for each region. At the time of analysis and interpretation, after looking at the overall results, one needs to evaluate sensitivity to various factors, including how each regional result contributed to the overall result. We will discuss approaches to determining sample size in each region and dealing with regional results.

Issues in Conducting Multiregional Trial: Experience in the PMDA Consultation Meetings

❖ Yuki Ando, Pharmaceuticals and Medical Devices Agency, 3 3 2 Kasumigaseki Chiyodaku, Tokyo, 100-0013 Japan, *ando-yuki@pmda.go.jp*

Key Words: multi-regional trial, simultaneous drug development

Recently an increasing number of Japanese pharmaceutical companies have planned to participate in multi-regional trials, in order to reach the goal of global simultaneous drug development and nearly simultaneous world-wide registration. Although it is desirable for such trials to be conducted under a common protocol among the participating regions, there may be a need to allow minor regional design differences e.g., concomitant medication. In this presentation, recent issues raised in PMDA consultation meetings that focused on global simultaneous drug development programs will be reviewed with a few examples in which regional protocol differences were needed.

17 Sufficient Dimension Reduction and High-dimensional Data ✪

IMS, Section on Nonparametric Statistics

Sunday, July 29, 2:00 pm–3:50 pm

Generalized K-Means Inverse Regression Estimation

❖ Xuerong Wen, University of Missouri-Rolla, 220 Rolla Bldg, 1870 Miner Circle, Rolla, MO 65409, *wenx@umr.edu*

Key Words: Dimension Reduction, Multivariate Regression, Central Subspace, K-means Clustering, Intra-cluster Information, IRE

Few methodologies are available for estimating the central subspace for regressions with multivariate responses due to the difficulties arising from slicing multivariate responses. Setodji and Cook(2004) introduced a new way of performing the slicing. They developed a method called k-means inverse regression (KIR), which makes use of the K-means algorithm to cluster the observed response vectors. However, their method ignored the intra-cluster information which could be substantial under some circumstances. In this paper, we proposed an improved method by incorporating the intra-cluster information into estimation. Our method outperformed KIR with respect to estimation accuracies of both the central subspace and its dimension. It also allows us to test the predictor effects in a model-free approach.

A Novel Moment-Based Sufficient Dimension Reduction Approach in Multivariate Regression

❖ Jae Keun Yoo, University of Louisville, 555 South Floyd Street, Dept of Statistics, Louisville, KY 40202, *peter.yoo@louisville.edu*

Key Words: sufficient dimension reduction, multivariate regression, moment-based method, predictor tests

Recently, a moment based dimension reduction methodology in multivariate regression, focusing on the first two moments, was introduced. We present in this paper a novel approach of the earlier method. This novel method possesses several desirable properties that the earlier method did not have such as dimension tests with chi-squared distributions, predictor effects test for the first two conditional moments without assuming any model, and so on. Simulated and real data examples are presented for studying various properties of the proposed method and for a numerical comparison of the presented method to the earlier method.

Principal Components for Regression: A Conditional Point of View

❖ Liliana Forzani, The University of Minnesota, 313 Ford Hall, School of Statistics, 224 Church Street S.E., Minneapolis, MN 55455, *liliana.forzani@gmail.com*

Key Words: Dimension Reduction, Principal component analysis

Many statistical applications involve regression with many predictors and problems arise when the number of predictors is large or they are very correlated. One way to tackle the problem is to reduce the dimension of the predictors. The focus of this talk is reduction of the dimension of the predictors from the Inverse Regression point of view. Assuming that the inverse regression X|Y follows a normal distribution with covariance independent of Y, Cook (2007) was able to find the linear combinations of the predictors X that are sufficient for the regression of X, in the sense that the distribution of Y|X is the same than the distribution of Y|fl’X for fl e Rpxd with d< p the smallest possible number. In this talk we present the maximum likelihood estimators for such combinations as well as testing procedures for this setting.

Dimension Reduction in Time Series

❖Jin-Hong Park, University of Georgia, 104 College Station Rd, E214, Athens, GA 30605, *jinpark@uga.edu*

Key Words: Time series central subspace, Kullback-Leibler distance, Density estimator, Nonlinear time series, Threshold

We develop a sufficient dimension reduction theory for time series, which does not require specification of a model but seeks to find a p times d matrix with the smallest possible number d such that the conditional distribution of current value given past vector with p lags is the same as that of current value given past linear combinations, resulting in no loss of information about the conditional distribution of the series given its past p values. We define the subspace spanned by the columns of a p times d matrix as the time series central subspace and estimate it using Kullback-Leibler distance. We show that the estimator is consistent. In addition, we propose a consistent estimate of d and a graphical method to determine the lag p. Finally, we present examples and real data analysis to illustrate the proposed theory, which may open new research avenues in time series data analysis.

Model-Free Multivariate Reduced-Rank Model Regression with Categorical Predictor

❖Messan C. Setodji, RAND Corporation, 1776 Main St, PO Box 2138, Santa Monica, CA 90407, *setodji@rand.org*; Lexin Li, North Carolina State University

Key Words: Dimension reduction, Central partial mean subspace, Multivariate regression, Multivariate reduced rank models

Cook and Setodji (2003) introduced the notion of model free reduced rank in multivariate regression when one is dealing exclusively with continuous predictors and developed a test of dimension for multivariate regression models. In this article, we propose an extension of the model free multivariate reduced-rank model to allow for a mixture of continuous and categorical predictors. The use of the method can be really useful at the onset of a multivariate analysis and an example of it's application to a study of the effect early life trajectories on aging is presented.

18 Spatial Statistics Applied to Environment ● ✪

IMS, Section on Statistical Consulting, Section on Statistics and the Environment

Sunday, July 29, 2:00 pm–3:50 pm

Block Bootstrap Procedures for Modeling Inhomogeneous Spatial Point Patterns

❖Yongtao Guan, Yale University, 60 College Street, New Haven, CT 06520-8034, *yongtao.guan@yale.edu*

Key Words: Block Bootstrap, Inhomogeneous Spatial Point Process

When modeling inhomogeneous spatial point patterns, it is often of interest to fit a parametric model for the first order intensity function (FOIF) of the process in terms of some measured environmental covariates. Estimates for the unknown coefficients can be obtained by maximizing a Poisson maximum likelihood criterion. We propose new block bootstrap procedures to estimate the variance of the estimated parameters. These procedures depend only on the FOIF of the process but not on any higher-order terms and thus can be easily applied once the FOIF has been estimated. We demonstrate the use of these procedures through both simulations and an application to a tropical forestry data example.

Assimilation of Current Measurements into a Circulation Model of Lake Michigan

❖Zepu Zhang, The University of Chicago, 5734 S. Ellis Ave, Chicago, IL 60637, *zpzhang@stanfordalumni.org*; Dmitry Beletsky, University of Michigan; David Schwab, National Oceanic and Atmospheric Administration; Michael Stein, The University of Chicago

Key Words: stream function, kriging, Princeton Ocean Model, data assimilation, hydrodynamic modeling

We present a method for assimilating current measurements into a 2-D circulation model of Lake Michigan. Measurements are assimilated by updating the stream function of the velocity field via kriging interpolation, therefore the physical constraint of a non-divergent flow is satisfied in the updated field. Moreover, the use of stream function avoids a component-wise treatment to the velocity vector. The coastal constraint is represented by the stream function being constant along the coastline, and is implemented by incorporating pseudo coastal data into the interpolation. This eliminates the need to construct complex spatial covariance models. The method also accommodates measurement uncertainties. Results show the method's ability to successfully meld measurements into simulations with the effect propagating in space and time, and the critical role of the covariance model.

Spectrum Estimation for Isotropic Intrinsically Stationary Spatial Processes

❖Chunfeng Huang, The Ohio State University, Department of Statistics, Columbus, OH 43210, *chuang@stat.ohio-state.edu*; Tailen Hsing, The Ohio State University; Noel Cressie, The Ohio State University

Key Words: Spectrum Estimation, Spatial Processes, Isotropic, Intrinsically Stationary, Regularized inverse problem

In the study of isotropic intrinsically stationary spatial processes, the spectral density function is a parameter of interest. A new spectrum estimation is formulated in terms of solving a regularized inverse problem. The regularized inverse problem is solved in a reproducing kernel Hilbert space as a constrained optimization problem. This methodology is applied to a spatial dataset of temperature changes over Americas.

Taxonomy of Spatially Clustered Disease Data: Capturing and Explaining Spatial Clustering and Cluster

❖Ge Lin, West Virginia University, 406 White Hall, Morgantown, 26506, *glin@wvu.edu*

Key Words: disease cluster, local, global, disease mapping, loglinear model

Spatial cluster detections often divided into general and focused tests. The general test is to detect a global or overall clustering trend, while the focus test is to identify the location of a local cluster. Although some tests are developed for the general, and some are developed for the focused, and some for both, the relationship between global and local has not been fully defined. In this paper, I first provide the taxonomy of global clustering and local clusters. I pointed out that global clustering and local clusters sometime are attributable the existence of local clusters, or the combination of both global and local trends. In some situations, even just a global trend could result the significant test results for both general and focused tests. Second, I define some global trends that may render different influence on the detection of a local cluster. Some global trends can be captured by spatial autoregressive models, some can be captured by explanatory variables, and some can be dealt with non-parametrically. Third, I relate global trends to common spatial events and point out that global trends are almost as a rule in spatial events rather than an exception. Both simulations and case studied are provided. In the case studied, infant mortality data from Guangxi province, China are used. It shows that without accounting for the global clustering trend, the detection of local clusters can be a moving target with significant local clusters almost covering the entire province. When the global trend is accounted for by elevation, which is a proxy of many factors, a number of significant local clusters can be detected.

Space-Time Latent Structure Modeling for Small-Area Health Data

❖ Andrew B. Lawson, University of South Carolina, Epidemiology Biostatistics, Columbia, SC 29209, *alawson@gwm.sc.edu*

Key Words: space-time, health, basis-functions, mixtures, MVCAR, latent

In this talk I will compare two basic approaches to space-time latent structure modeling. The first approach will be focused on the use of mixtures where a weighted linear combination of m temporal profiles is assumed. The temporal components have temporal dependence, flexibly defined by autoregressive priors. The weight components are scaled to yield probabilities over the components but are allowed to have cross-correlation and spatial correlation between sites. Prior comparison of these approaches is made to some county-level lung cancer mortastributions are assumed for both the temporal and spatial components. A second approach to this form of modeling is the use of linear combinations of basis functions where the regression parameters have spatially structured prior distributions. These models allow the use of zero-inflated mixture priors to let components drop from the model.

19 Bayesian Modeling of Longitudinal or Event-Time Data Subject To Nonresponse or Censoring ● ✪

Section on Bayesian Statistical Science, ENAR, Section on Health Policy Statistics

Sunday, July 29, 2:00 pm–3:50 pm

Bayesian Variable Selection in Linear Mixed-Effects Models for Longitudinal Data with Missing Values

❖ Xiaowei Yang, University of California, Davis, 2221 Caravaggio Drive, Davis, CA 95618, *xdyang@ucdavis.edu*

Key Words: Bayesian Variable Selection, Multiple Imputation, Mixed-Effects Models, Missing Data

Mixed-effects models are popularly applied in statistical practice with longitudinal data for which a challenge arises from the missing values on either outcome or predictor variables. For the purposes of model selection, we propose two imputation-based strategies: ITS (impute, then select) and SIAS (simultaneously impute and select). ITS applies the method of multiple imputation directly to first generate several complete data sets, then conduct Bayesian variable selection on each imputed data set, and finally combine the selection results. SIAS embeds the data imputation and variable selection steps within one integrated Gibbs sampler. By applying the two strategies with various Bayesian variable selection algorithms to both practical and simulated data sets, we evaluate their performance and compare them with traditional methods (step-wise, AIC, BIC, etc.).

Bayesian Variable Selection for Analyzing Longitudinal Substance Abuse Treatment Data Subject to Informative Censoring

❖ Susan Paddock, RAND Corporation, 1776 Main Street, Santa Monica, CA 90401, *paddock@rand.org*

Key Words: Bayesian variable selection, conditional linear model, informative censoring, longitudinal data, nonignorable nonresponse, pattern-mixture model

Measuring the process of care in substance abuse treatment requires analyzing repeated client assessments at critical time points during treatment tenure. These assessments are frequently censored due to early departure from treatment. Informative censoring is often characterized by the last observed assessment time. However, if missing assessments for those who remain in treatment are attributable to logistical reasons rather than to the underlying treatment process being measured, then length of stay might better characterize censoring than would time of measurement. In this talk, I will describe how to incorporate Bayesian variable selection into the Conditional Linear Model to assess whether time of measurement or length of stay better characterizes informative censoring while incorporating uncertainty about the effect of censoring on treatment process change into the analysis.

Approximate Bayesian Model Averaging for Latent Class Pattern Mixture Models with an Unknown Number of Classes

❖ Jason Roy, Geisinger Health System, Center for Health Research, MC:44-00, 100 N. Academy Ave, Danville, PA 17822, *jaroy@geisinger.edu*; Michael Daniels, University of Florida

Key Words: missing data, dropout, latent variable, longitudinal data

We consider the problem of fitting pattern mixture models to longitudinal data when there are many unique dropout times. We propose a marginally specified latent class pattern mixture model. The marginal mean is assumed to follow a generalized linear model, while the mean conditional on the latent class and random effects is specified separately. Because the dimension of the parameter vector of interest (the marginal regression coefficients) does not depend on the assumed number of latent classes, we propose to treat the number of latent classes as a random variable. We specify a prior distribution for the number of classes, and calculate (approximate) posterior model probabilities. In order to avoid the complications with implementing a fully Bayesian model, we propose a simple approximation to these posterior probabilities.

Institutionalization at Follow-Up: An Application of Principal Stratification to the Evaluation of Substance Abuse Treatment

❖ Daniel McCaffrey, RAND Corporation, 4570 Fifth Avenue Suite 600, Pittsburgh, PA 15213-2665, *danielm@rand.org*; Beth Ann Griffin, RAND Corporation

Key Words: potential outcomes, treatment effects

Participants in longitudinal studies on the effects drug treatment and criminal justice system interventions are at high risk for institutionalization (i.e., spending time in an environment where their freedom to use drugs, commit crimes, or engage in risky behavior may be circumscribed). Methods used for estimating treatment effects in the presence of institutionalization during the follow-up can be highly sensitive to assumptions that are unlikely to be met in applications. In this paper, we extend the methods of principal stratification to model institutionalization at follow-up for the estimation of the effect of residential substance abuse treatment versus outpatient services. We discuss practical issues in applying the model to data and show via simulation that the model can recover true effects, but that careful selection of starting values is essential for consistent estimation.

Latent Class Modeling of Growth and Reproduction in Sublethal Toxicity Tests

❖Michael Pennell, The Ohio State University, 320 West 10th Avenue, Columbus, OH 43210, *mpennell@sph.osu.edu*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Bayesian methodology, Dirichlet process, Latent variables, Mixture model, Order restrictions, Toxicology

Medium-throughput methods have been established to rapidly screen many developmental toxins using Caenorhabditis elegans. In these studies, organisms are exposed to one of 12 doses and their size is measured after 72 hours, during which time unaffected organisms may also reproduce. This presents some interesting statistical challenges as 1.) there is no way of distinguishing the progeny from the original experimental units and 2.) there is no consensus on how to define larval stage in terms of size. To address these issues, we propose a latent class model for growth and reproduction. We model the number of organisms at each dose using a Dirichlet process mixture of Poissons and size using a mixture of two multivariate t distributions. We also use mixture priors to test for changes across dose with respect to growth and reproduction and apply order restrictions to increase efficiency.

20 Section on Bayesian Statistics Student Paper Competition ● ✪

Section on Bayesian Statistical Science

Sunday, July 29, 2:00 pm–3:50 pm

Bayesian Classification of Sleep States in Mice (Student Paper Competition)

❖Chaitra Nagaraja, University of Pennsylvania, 2400 Chestnut St, Wharton School, Apt 3212, Philadelphia, PA 19103, *chaitra@wharton.upenn.edu*; Shane T. Jensen, University of Pennsylvania; Abraham Wyner, University of Pennsylvania

Key Words: Bayesian, classification, hidden Markov model, Viterbi

We examine the problem of inferring sleep states (awake, non-REM sleep, and REM sleep) in mice based on video data. Observed measures of velocity, size, and aspect ratio of a mouse were taken from video images over extended time periods. Data from one mouse was used to fit a hidden Markov model with the sleep states for each observation as latent variables. Mixture modeling in the emission densities was needed to account for multi-modality in the observed variables. We compare the accuracy of predictions based on the optimal path of sleep states generated from a Viterbi algorithm with predictions generated directly with the posterior state probabilities for each observation. The proposed Bayesian procedure provides more accurate classifications.

Path Sampling To Compute Bayes Factors: An Adaptive Approach

❖Genevieve Lefebvre, McGill University, 805 Sherbrooke W, Burnside Hall room 1005, Montreal, QC H3A 2K6 Canada, *lefebvre@math.mcgill.ca*; Russell Steele, McGill University; Alain C. Vandal, McGill University; Sridar Narayanan, McConnell Brain Imaging Centre; Douglas L. Arnold, McConnell Brain Imaging Centre

Key Words: Model Selection, Monte Carlo Integration, Adaptive Quadrature, Mixed-Effects Model, Importance Density

Performing model selection using Bayes factors (BF) is a challenging task, particularly when the models are large and complex. Path sampling (PS) is recognized as one of the most powerful Monte Carlo integration methods for BF estimation. We examine the impact of two tuning parameters of PS, the specification of the importance density and of the grid, which are shown to be potentially very influential. We then propose the use of an algorithm to automate the selection of the grid in PS, the Grid Selection by Adaptive Quadrature (GSAQ) approach. A bound for the bias of the corresponding PS estimator is also provided. We perform a comparison between GSAQ and standard grid implementation of PS using a well-studied dataset; GSAQ is found to yield superior results. GSAQ is then successfully applied to longitudinal regression models in Multiple Sclerosis research.

Multivariate Developmental Trajectory Models To Identify Longitudinal Patterns of Disability in Seniors

❖Jason Connor, Berry Consultants, 117 Westchester Blvd, Noblesville, IN 46062, *jason@berryconsultants.com*

Key Words: Bayesian analysis, Mixture models, Longitudinal data, Health policy, Disability

I introduce two analytical methods which extend group-based trajectory models to multivariate outcomes. The first, the marginal developmental trajectory model, identifies latent patterns separately for each outcome, then uses a contingency table framework to identify common combinations of trajectories. The second, the joint developmental trajectory model, considers all outcomes simultaneously to identify combinations of trajectories. I use group-based finite mixture models to identify and describe latent subpopulations for the multivariate outcomes of interest. I apply these methods using data from the National Long Term Care Survey, which measures disabilities in American elderly from 1982 to 2004. Within latent classes, the models clearly illustrate the age of onset and order in which disability patterns typically appear. Finally, I describe how disability is changing over time.

Analysis of Immunohistochemical Data Using Measurement Error Models

❖Ronglai Shen, University of Michigan, 1789 Beal Ave Apt 2, Ann Arbor, MI 48105, *rlshen@umich.edu*; Debashis Ghosh, University of Michigan; Jeremy Taylor, University of Michigan

Key Words: Biomarker, Immunohistochemical data, Latent Expression Index, Measurement error, Tissue Microarray

The advent of Tissue Microarrays (TMAs) has provided a proteomic platform for validating cancer biomarkers emerging from large-scale DNA microarray studies. Repeated observations from each tumor result in substantial biological and experimental variability. We propose to analyze TMA data with patient survival endpoints in a measurement error model framework. Two goals are explored: 1) in a two-stage approach, a Latent Expression Index (LEI) is introduced as a summary protein expression index estimated from the TMA measurements; 2) a joint model of survival and TMA expression data is established via a shared random effect. Bayesian estimation is carried out using a Markov Chain Monte Carlo (MCMC) method. In a case study, we compared the expression estimates from the proposed error methods in differentiating risks of developing PSA failure among surgically treated prostate cancer patients.

21 Identifying and Overcoming Barriers to Teaching a Reformed Introductory Course

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences
Sunday, July 29, 2:00 pm–3:50 pm

Identifying and Overcoming Barriers to Teaching a Reformed Introductory Statistics Course

❖Marjorie Bond, Monmouth College, 218 E. Girard Avenue, Monmouth, IL 61462, *mebond@monm.edu*; ❖Christine Kohnen, Saint Olaf College, 1520 Saint Olaf Ave, MSCS, Northfield, MN 55124, *kohnen@stolaf.edu*; ❖Karen Kinard, Tallahassee Community College, 3440 Thresher Drive, Tallahassee, FL 32312, *kinardk@tcc.fl.edu*; ❖Mary Parker, Austin Community College, 1622 Waterston Ave, Austin, TX 78703, *parker@math.utexas.edu*

Key Words: time, change, technology, tenure, interpersonal, intrapersonal

Much effort has been expended to develop materials and methods to be used in teaching reformed introductory statistics courses. However, there remain substantial barriers to their production and implementation. These barriers include time, change, and technology. Interestingly enough, each of these is a barrier in a different way for faculty, students, and administrators involved in these courses. During this session we will present questions we have identified relating to these barriers, solicit additional questions from the audience, do a Pareto analysis to identify the audience's preferences, and address as many questions as possible, in order of the audience's preferences.

22 Bayesian Time Series and Binary Data ● ✪

Section on Bayesian Statistical Science
Sunday, July 29, 2:00 pm–3:50 pm

Mixtures of State Space Models for Time Series with Structural Breaks

❖Ori Rosen, University of Texas at El Paso, Department of Mathematical Sciences, 221 Bell Hall, El Paso, TX 79968, *ori@math.utep.edu*; Sally Wood, University of New South Wales; Robert Kohn, University of New South Wales

Key Words: State-space Models, Kalman Filter, Mixtures-of-Experts, MCMC

We consider mixtures of time series models which can be expressed in state space form. The mixing weights depend on time and on unknown parameters resulting in a flexible model that can accommodate varying local behavior of the time series. The number of components is treated as one of the model parameters. An example is a mixture of autoregressive models in which case both the common lag and the number of components are assumed unknown. We take a Bayesian approach and sample all the parameters from their posterior distribution. The methodology is illustrated with simulated data, as well as with real data.

Bayesian Analysis and Applications of a Generalized Threshold Autoregressive Model

❖Peng Sun, Merck & Co., Inc., PO Box 1000 UG1D44, North Wales, PA 19454, *peng_sun@merck.com*

Key Words: Bayesian, Gibbs Sampler, Threshold Autoregressive Model

In this paper, we generalize the self exciting threshold autoregressive (SETAR) model by allowing the threshold variable to be a linear combination of functions of lag values of the series. The generalized model is then analyzed in the Bayesian framework. When applying the model to simulated datasets, we find that the performance of the posterior simulator is highly dependent on the choice of starting values. If the starting value is close to the true parameter value, then the chain can easily achieve convergence. Otherwise, it persists near a local mode of the posterior distribution. The convergence problem is explored in detail and an effective sequential algorithm is developed to address this issue. We present an application of the generalized model to the S&P 500 daily return data. The model fitting results compare favorably to the GARCH(1,1) model fitted to the same dataset.

Bayesian Clustering of Time Series with Change Points

❖Garrick Wallstrom, University of Pittsburgh, Suite M-183 Parkvale Building, 200 Meyran Avenue, Pittsburgh, PA 15260, *garrick@cbmi.pitt.edu*; William Hogan, University of Pittsburgh

Key Words: Clustering, Bayesian, Time Series, Change point, Biosurveillance

A common problem in biosurveillance is selecting time series to aggregate and monitor for the detection of disease outbreaks. Selection of time series is complicated by qualitative changes in the time series that occur at unknown points in time. We present a hierarchical Bayesian model for clustering time series with change points. We use Markov chain Monte Carlo simulation to estimate posterior distributions and point estimates for model parameters. We also estimate smooth curves that represent typical behavior for the time series within each cluster. We apply the method to construct aggregates of product-level sales data for over-the-counter health care products.

Estimating Kappa Coefficient and Tetrachoric Correlation for Clustered Binary Data

❖Xiao Zhang, The University of Alabama at Birmingham, 1665 University Blvd. 414A, Birmingham, AL 35294, *xzhang@ms.soph.uab.edu*; Gary Cutter, The University of Alabama at Birmingham

Key Words: Kappa Coefficient, Tetrachoric Correlation, Multivariate Probit Model, Markov chain Monte Carlo, Simulation Study, Sensitivity

We estimate the kappa coefficient and the tetrachoric correlation for clustered binary outcomes from a Bayesian perspective. We construct a Bayesian model and develop a Markov chain Monte Carlo (MCMC) algorithm to sample the tetrachoric correlation. Based on the definition of the kappa coefficient which is the function of the mean and the tetrachoric correlation of the binary outcomes, our method produces the posterior inference of the kappa coefficient. We investigate the sensitivity of the posterior estimation of the kappa coefficient and the tetrachoric correlation. We illustrate our method through the psychiatric data used by Lipsitz, Laird, and Brennan (1994) and through the data from an epitope-specific immunotherapy in rheumatoid arthritis study.

Selected Bayesian Inference for Correlated Binary Data

❖Patricia Bahnsen, Stephen F. Austin State University, SFA Station Box 13040, Nacogdoches, TX 75962, *pbahnsen@sfasu.edu*

Key Words: Bayesian methods, correlated binary data, structural zero model, bivariate binomial model

Correlated binary measurements can occur in a variety of practical contexts and afford interesting statistical modeling challenges. In order to model the separate probabilities for each measurement we must somehow account for the relationship between them. Here we consider Bayesian models for three different prior structures using likelihoods representing the data in the form of a 2 ? 2 table. To do so, consider the data as counts resulting from two correlated binary measurements: the onset of diabetic retinopathy and the onset of diabetic nephropathy. We compare resulting posterior distributions from a Jeffreys' prior, independent beta priors, and conditional beta priors, based on a structural zero likelihood model and the bivariate binomial model.

Bayesian Inference in 2 by 2 Tables

❖Yong Chen, Johns Hopkins Bloomberg School of Public Health, 4014 Linkwood Rd, Apt F, Baltimore, MD 21210, *yonchen@jhsph.edu*; Sining Chen, Johns Hopkins Bloomberg School of Public Health; Haitao Chu, Johns Hopkins Bloomberg School of Public Health

Key Words: Bayesian Inference, 2 by 2 table, Dirichlet priors, exact distribution, Relative risk, Risk difference

The relative risk, odds ratio and risk difference are the most commonly used measures of the association between binary exposure and outcome. Using conjugate Dirichlet priors and Beta priors, we derive the exact posterior distributions of these measures. Bayesian inference based on these exact distributions is discussed, particularly when the sample size is small. The proposed method is applied to several biomedical studies.

Classification with Spectroscopic Data Using Bayesian Variable Selection

❖Hongxiao Zhu, Rice University, 6100 Main St MS 138, Houston, TX 77005, *hxzhu@stat.rice.edu*; Dennis D. Cox, Rice University

Key Words: Bayesian Variable Selection, Classification, Probit Model, Latent Variable, Fluorescence Spectroscopy, Cervical Cancer

Fluorescence Spectroscopy provides a non-invasive tool for in vivo, real time diagnosis of cervical pre-cancer. An important issue involved is to accurately classify diseased tissue from normal using extremely high dimensional data-the Fluorescence excitation-emission matrices (EEMs). This paper presents a Bayesian probit model for classification purpose based on features extracted from EEMs. Latent variables are included to simplify computation, and variable selection is performed through a mixture normal prior of the regression coefficients. MCMC methods-a Gibbs sampler and a Hybrid Metropolis-Hasting/Gibbs sampler, are used to obtain posterior samples. Simulations show that this model produces accurate variable selection results and good estimates of the regression coefficients. Application to spectroscopic data gives improved classification performance compared with other methods.

23 Recent Advances in Microarray Data Analysis IV ● ✪

Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

A Statistical Method for Selecting Differentially Expressed Genes

❖Zhao Chen, Florida Gulf Coast University, 10501 FGCU Blvd South, College of Arts and Sciences, Fort Myers, FL 33965, *kchen@fgcu.edu*

Key Words: microarray, differentially expressed genes

Typical Microarray data has tens of thousands of random variables and only several samples. When we graph a scatter plot of the ratio over intensity, it has pear shape, which means more variation at the lower intensity and less variation at the higher intensity. Due to this unusual property, this paper provides a statistical method for selecting differentially expressed genes

Inferring Genomic Copy Number from CGH Arrays

❖Stephen Erickson, The University of Alabama at Birmingham, 327 Ryals Public Health Building, 1665 University Boulevard, Birmingham, AL 35233, *erickson@uab.edu*

Key Words: empirical Bayes, copy number polymorphism, comparative genomic hybridization

There is growing evidence that large-scale copy number polymorphisms (CNPs) contribute significantly more to human genetic variation and diversity than recently thought. Variations in CNPs have been implicated in a variety of disease phenotypes, especially cancers. Comparative genomic hybridization (CGH) arrays simultaneously assay tens of thousands of loci and are used to infer genome-wide copy number profiles. Like all genomic microarray tiling technologies, however, CGH arrays are subject to measurement error and other inconsistencies. We propose an empirical Bayes approach to infer copy number profiles from CGH arrays. The model structure enables the sharing of power between loci, particularly in estimating measurement error, and incorporates a spatial structure to facilitate the identification of CNPs which span multiple CGH loci.

A Ridge Penalized Principal-Components Approach Based on Heritability for High-Dimensional Data

❖Yuanjia Wang, Columbia University, 1531 9th St, Fort Lee, NJ 07024, *yw2016@columbia.edu*; Yixin Fang, Columbia University; Man Jin, Columbia University

Key Words: Principal-component analysis, Ridge penalty, Heritability, Family data, Cross validation, High-dimensional data

Microarray technique allows measurement of thousands of gene expression levels simultaneously and provides opportunity for mapping shared genetic contribution to multiple expression levels. Clustering analysis and principal component analysis are proposed to reduce the dimensionality of phenotypes. Genetic linkage or association analysis is then applied to the combined phenotypes. However, the usual clustering and principal component analysis are only appropriate for independent data. In most of the genetic studies where family data are available, applying these standard analyses to founders leads to loss of information. Here are proposed a clustering and a principal component approach based on heritability for high-dimensional data that take into account of the family structure information. The methods are illustrated through an application to human lymphoblastoid gene expression data.

Robustness of Cancer Risk Prediction Using Microarrays

❖Chiang-Ching Huang, Northwestern University, 680 N. Lake Shore Dr., Suite 1102, Chicago, IL 60611, *huangcc@northwestern.edu*; Irene Helenowski, Northwestern University

Key Words: cross validation, cancer, risk stratification, sample size, microarray

Microarrays have demonstrated its utility in cancer research for better risk stratification. However, classifiers constructed from several cancer studies cannot be confirmed in subsequent studies, raising the concern regarding the robustness of cancer classification using microarrays. This hurdle is largely due to the limited number of tumors under study and cancer heterogeneity evidenced by the gene expression profiles. We propose to use a sequential cross validation approach to investigate the sample size required to achieve statistical robustness of a molecular classifier. This approach also serves to identify tumors that can not be correctly classified by the gene expression profiles of tumor tissues alone. Our result suggests current sample size calculation for microarray studies, especially in the prognosis problem, could be underestimated.

Incorporating Prior Knowledge of Predictors into Penalized Classifiers with Multiple Penalty Terms

❖ Feng Tai, The University of Minnesota, 1979 Knapp St, Apt W4, St Paul, MN 55108, *fengtai@biostat.umn.edu*; Wei Pan, The University of Minnesota

Key Words: gene function, PAM, PPLS, penalized classifier

In the context of sample classifications with microarray data, many methods have been proposed. However, almost all the methods ignore existing biological knowledge and treat all the genes equally a priori. On the other hand, because some genes have been identified by previous studies to have biological functions or to be involved in pathways related to the outcome, incorporating this type of prior knowledge into a classifier can potentially improve both the predictive performance and interpretability of the resulting model. We propose a simple and general framework to incorporate such prior knowledge into building a penalized classifier. We group the genes according to their functional associations based on existing biological knowledge or data, and adopt group-specific penalty terms and penalization parameters. We apply the idea to PAM and PPLS and show the improvement.

Optimal Microarray Designs for cDNA Microarray Experiments

❖ Tathagata Bandyopadhyay, Indian Institute of Management Ahmedabad, Wing 14 Indian Institute of Management A, Vastrapur, Ahmedabad, 380015 India, *tathagata@iimahd.ernet.in*

Key Words: microarray experiment, factorial experiment, optimal design

We investigate optimal factorial designs for cDNA microarray experiments, where the parameters of interest differ from traditional factorials. We apply approximate theory to obtain analytical results for 22 factorial. More general, including asymmetric factorials are considered next and tools of Kronecker representation and unimodularity are employed to derive strong exact result on optimal saturated designs. Its impact on the study of optimal or efficient nearly saturated designs is also investigated. These results are then extended to more intricate situation where the underlying model incorporates dye-coloring effects, and finally the role of dye-swapping is critically examined.

An Integrated Approach Using Nonparametric Local Correlation to Identify Critical Pathways in Disease

❖ Yian Chen, Texas A&M University, 437 Blocker Building, Department of Statistics, College Station, TX 77843, *annchen@stat.tamu.edu*; Marina Vannucci, Texas A&M University

Key Words: nonlinearity, local correlation, data integration, microarray, pathway

Nonlinearity in high-dimensional biology data and their integration are challenging problems in the post-genomic era. Inspired by the correlation integral analysis used in physics to measure fractal dimensions, we propose a nonparametric metric, maximum local correlation (MLC), to detect nonlinear association. A delta statistic based on MLC is proposed to detect critical genes and pathways by measuring association change between the control and experimental animals. Data integration is performed through extracting principal components (PCs) of genes using expression data for each pathway, and then detecting association changes among the PCs between the control and experiment. We applied our method on a microarray dataset from the control mouse and the SWV mouse, which is susceptible to neural tube defects, induced by hyperthermia. The results include promising candidate pathways.

24 Recent Advances in Longitudinal Data Analysis ● ✪

Biometrics Section, Biopharmaceutical Section

Sunday, July 29, 2:00 pm–3:50 pm

Self-Reported Behavior of High-Risk HIV+ Youth: Random Effects Models for Multivariate Longitudinal Discrete Data

❖ Lily Altstein, University of California, Los Angeles, 825 Weyburn Terrace, C20, Los Angeles, CA 90024, *laltstein@ucla.edu*

Key Words: Multivariate categorical data, Bayesian methods, Behavioral intervention, Sero-status disclosure

This analysis applies generalized linear mixed models to analyze longitudinal data from an interventional study encouraging HIV positive adolescents to attenuate high risk sexual behavior. Multivariate outcomes of interest include number of sex partners, type of partner, and the frequency at which the subject disclosed his HIV status. Bayesian methods are used to estimate the effect of the intervention on safe sex behavior over time.

Longitudinal Analysis for Post-Transplant Serum Creatinine Data

❖ Zailong Wang, Novartis Pharmaceuticals, 50 Ridge Drive, Livingston, NJ 07039, *zailong.wang@yahoo.com*; Luen Lee, Novartis Pharmaceuticals

Key Words: mixed-effect regression, covariance structure analysis, phi-divergence statistic, goodness-of-fit test, treatment effect, serum creatinine data

In transplant clinical studies, serum creatinine (SCr) are measured longitudinally on the same individuals to monitor and assess immunosuppressive effects on renal function. Hence, longitudinal analysis methods should be appropriate to model the SCr data for identifying predictors among immunosuppressive regimens and potential covariates of renal function by testing the corresponding hypotheses. In this paper, we propose the use of mixed-effect regression model to analyze SCr data that were collected from a Novartis transplant clinical trial. The analysis covers the selection of fixed-effect factors and variance-covariance matrix structure, and testing on treatment effects. The residuals from the chosen "best" model are analyzed for goodness-of-fit test using phi-divergence statistic which is a generalized likelihood ratio or Pearson Chi-square statistic.

Giving Treatment to Controls: When Is It a Good Idea?

❖ Kimberly Walters, The Ohio State University, Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *walters@stat.ohio-state.edu*

Key Words: linear mixed effects, longitudinal methodology, simulation studies, treatment estimation

The power of the likelihood ratio test to detect treatment effect was compared for a full linear mixed effects model and a proposed reduction. Simulation studies were conducted under conditions of equal or unequal slopes in the treatment and control groups following intervention. Initial parameter values were generated by fitting the full model to data from a study in which all participants received treatment, either upon entry or following a control period. The full model does not assume equal post-intervention slopes while the reduced model does, therein combining information from both groups to determine the effect of intervention. We anticipate the reduced model will offer more power to detect treatment effect and an improved estimator of that effect when the post-intervention slopes are not too dissimilar.

Comparing the Correlation Coefficients Between the Random Effects in Multivariate Linear Mixed Models

❖ Yuxiao Tang, Rush University, 903 S Ashland Ave Apt 612, Chicago, IL 60607, *yuxiao_tang@rush.edu*; Chengjie Xiong, Washington University in St. Louis

Key Words: linear mixed models, longitudinal study, correlation coefficients, random effects

As the linear mixed models have been extensively employed in longitudinal studies, the demands for the appropriate methods for the inferences concerning the correlations among random effects have increased. This paper considers the situation when multiple outcome measures are longitudinally observed from the same subjects under study and the study objective is to compare the strength of the correlations among the rates of longitudinal changes from these outcome measures. Under the framework of multivariate random intercept and random slope models for longitudinal data, we propose two methods for testing equality of the correlation coefficients between random slopes from multiple outcome measures: one is based on Fisher's z-transformation and the other is based on likelihood ratio test. Simulations are performed to compare the powers of the two tests.

Comparison of ICC and CCC for Assessing Agreement for Data Without and with Replications

❖ Chia-Cheng Chen, North Carolina State University, 2320 Apt 308, Crescent Creek Dr, Raleigh, NC 27606, *cchen4@ncsu.edu*

Key Words: Accuracy, Precision, Agreement, Reliability, Intraclass Correlation Coefficient, Concordance Correlation Coefficient

ICC has been traditionally used for assessing reliability between multiple observers for data with or without replications. Definitions of different ICCs depend on the assumptions of ANOVA models. The parameter estimator for the ICC is based on the method of moment with the underlying assumed ANOVA model. However, often times the ANOVA assumptions are not met in practice and researchers may compute these estimates without verifying the assumptions. It is not clear what these estimators are estimating if the ANOVA assumptions are not met. We compute the expected value of the ICC estimator under a very general model to get a sense of the population parameter that the ICC estimator provides. We compare this expected value to another popular agreement index, CCC. The main findings are reported for data without and with replications for three versions of ICCs.

A Local Sensitivity Analysis Approach to Longitudinal Binary Data with Nonignorable Dropout

❖ Hui Xie, University of Illinois at Chicago, 1924 Wilmette Avenue, Unit C, Wilmette, IL 60091, *huixie@uic.edu*

Key Words: Missing data, Missing not at Random, Nonignorability

Longitudinal binary data subject to potentially nonignorable dropout is a challenging problem. Frequently an analysis has to rely on some strong but unverifiable assumptions, among which ignorability is a key one. Sensitivity analysis has been advocated to assess the likely effect of alternative assumptions about dropout mechanism on such an analysis. Previously Ma et al. (2005) applied a general index of local sensitivity to nonignorability (ISNI} (Troxel et al. 2004) to measure the sensitivity of MAR estimates to small departures from ignorability for continuous longitudinal outcomes. We extend the ISNI methodology to handle longitudinal binary data subject to nonignorable dropout. Through a simulation study, we evaluate the performance of the proposed methodology. We then illustrate the method in two real examples.

A Comparison of Mixed Effect Models and Cox Proportional Hazard Models for Unbalanced Longitudinal Data

❖ Wonsuk Yoo, Wayne State University, 4201 St Antoine St UHC 4H, Detroit, MI 48201, *wyoo@med.wayne.edu*; Dennis Tsilimingras, Wayne State University; John M. Flack, Wayne State University

Key Words: longitudinal data, unbalanced intervals, mixed effect model, Cox proportional hazard model, time dependent covariates, MedTrace

There has been huge development in research of hypertension and cardiovascular disease area during last ten years. Most research shows that they provided clinical and epidemiologic evidence between time-dependent blood pressure levels and various study subjects of interest. Even though there has been much research including various endpoint and analysis methods, many studies had used two statistical models including mixed effect models and Cox proportional hazard models with different risk factors as their covariates. This research compare mixed effect models and Cox proportional hazard models with important differences of statistical results due to association among risk factors, unbalanced intervals between visits, and different endpoints. We use MEDTRACE (The Medical Education, Tracking, Reporting and Clinical Evaluation) data, a HIPAA compliant electronic health record.

25 Statistical Issues in Discovery and Nonclinical Development ● ✪

Biopharmaceutical Section, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

A New, Convolution-Based Approach to Developing a Level-A In Vitro In Vivo Correlation

❖Roel Straetemans, Johnson & Johnson PRD, Turnhoutseweg 30, Beerse, 2340 Belgium, *rstraete@prdbe.jnj.com*; Tom Jacobs, University Hasselt; Geert Molenberghs, Hasselt University; Luc Bijnens, Johnson & Johnson PRD

Key Words: IVIVC, one-stage, convolution, modeling

An IVIVC model describes the relationship between an in vitro property of a dosage form and a relevant in vivo response. IVIVC models find multiple practical applications within the pharmaceutical research, e.g. in reducing the number of human studies during the clinical development (CDER, 1997). Initial work on IVIVC modeling was done two decades ago (Gillespie et al., 1985) with the introduction of the so-called deconvolution method. This method, however, has some drawbacks and O'Hara et al. (2001) developed an alternative, based on a convolution procedure and involving a two-stage analysis. We developed a convolution-based procedure, novel in involving a one-stage approach and allowing mixture distributions for heterogeneous formulations (Jacobs et al., 2006). The method is successfully applied to a Galantamine controlled release formulation, a treatment for Alzheimer's disease.

Application of Group Sequential Methodology to a Preclinical Model

❖ Lixin Han, Wyeth, 35 Cambridgepark Drive, MS 35 3100 60, Cambridge, MA 02140, *hanl2@wyeth.com*; Suzana Marusic, Wyeth

Key Words: preclinical, group sequential, Type I error, power, simulation

Group sequential methods have been used in clinical trials extensively and some have advocated their use in preclinical studies. In this presentation, multiple group sequential methods are applied to a preclinical model, in which multiple compounds are compared to a control in repeated experiments for a quantitative response. Currently, the statistical approach to this preclinical model is to analyze the data by analysis of variance after data of each experiment are collected; the results from repeated experiments are combined by applying a decision rule without formally controlling for overall Type I error rate due to repeated testing. The group sequential methods are compared to the current approach in terms of Type I error, power, benefits and drawbacks.

Developing Statistical Methods for Setting Specification Limits

❖Michelle Quinlan, University of Nebraska-Lincoln, 1601 N 35th St, Lincoln, NE 68503, *mquinlan22@yahoo.com*; Erin Blankenship, University of Nebraska-Lincoln; Walt Stroup, University of Nebraska-Lincoln; James Schwenke, Boehringer Ingelheim Pharmaceuticals, Inc.

Key Words: shelf life, specification limit, confidence/prediction interval, random effect

A product's shelf life can be estimated from the time of first intersection between the specification limit for a stability limiting characteristic of a pharmaceutical product and an agreed upon interval estimate. In addition to defining a shelf life, a specification limit also defines bounds for future batches to confirm a product's shelf life. Any statistical methodology developed to facilitate setting specification limits must provide a reliable estimate of future batches. In addition, when multiple batches are used to support setting specification limits, controversy exists as to whether batches should be considered fixed or random effects. After assessing alternative interval estimate approaches, the goal is to develop multi-batch (fixed/random) extensions and compare the alternatives via simulation to determine which yield interval estimates suitable to shelf life estimation.

Statistical Consideration in the Evaluation of Red Blood Cell Products

❖Jeongsook Kim, Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852, *jessica.kim@fda.hhs.gov*; Tie-Hua Ng, Food and Drug Administration

Key Words: in vivo RBC survival, confidence interval, population proportion of successes, radiolabeled RBC

This presentation discusses the acceptance criteria for autologous 24-hour radiolabeled red blood cell in vivo survival studies in the evaluation of RBC products. The current acceptance criteria is based on the individual unit RBC survival, that is a one-sided 95% lower confidence limit for the population proportion of successes >70%, where a "success" is defined as an individual unit whose in vivo RBC survival is at least 75%. Alternative approaches to meeting this acceptance criterion will be discussed.

Optimization of Dilution Scheme

❖Andrew Chen, MedImmune, Inc., One Medimmune Way, Gaithersburg, MD 20878, *chenxi@medimmune.com*; Harry Yang, MedImmune, Inc.; Iksung Cho, MedImmune, Inc.

Key Words: ELISA assay, Serial dilution

The ELISA is a fundamental tool of clinical immunology, and is used both as screening test and potency assay. Based on the principle of antibody-antibody interaction, this test allows for easy visualization of results and can be completed without the additional concern of radioactive materials use. However, it is also well known that the ELISA assay is, in general, very variable. There are many sources of variations such as day, analytical, reagent and etc. that may contribute to the total variability. In this paper, through statistical modeling, we develop a method that allows us to choose an optimal dilution scheme for sample processing so as to reduce the overall assay variability. The theoretical results were verified through a simulation study.

Statistical Analysis of Measurement Data with Values Below Detection Limits

❖Donghui Zhang, sanofi-aventis, 1041 Rt 202 206 PO Box 6800, BRW M203A, Bridgewater, NJ 08807, *donghui.zhang@sanofi-aventis.com*; Cun-Hui Zhang, Rutgers University

Analytical laboratory measurements are often reported as less than detection limits (LDL). This poses challenges for statistical analysis unless the percentage of data below LDL is small. This talk will discuss several statistical approaches.

One-Sided Test for Gene-Gene Association

❖Chaehyung Ahn, Rho, Inc., 4033 Remington Oaks Circle, Cary, NC 27519, *cahn@rhoworld.com*

Key Words: one-sided test, contingency table, mutation, consensus, exact method

The genetic code can be considered as nominal data, since there is no known way to quantify each character. Therefore, when we evaluate an association between two positions, we usually rely on test statistics derived from a contingency table that asymptotically follow a chi^2-distribution. Such test statistics are by nature two-sided. To develop a one-sided method, we note that genes tend to be very conservative, and as a result, viable mutations are so rare that usually only one dominant character (the consensus) is allowed at each position. Exploiting this fact, we develop a one-sided test to evaluate the association with respect to the direction away from the consensus. Appropriate test statistics are presented with an exact method for assessment of statistical significance.

26 Generalized Distributions and Testing Methods

General Methodology, Biopharmaceutical Section, IMS, Section on Physical and Engineering Sciences, SSC

Sunday, July 29, 2:00 pm–3:50 pm

Iterative Modified Likelihood Ratio Test for Homogeneity

❖Pengfei Li, University of Waterloo, Unit 305, 350 Columbia Street West, Waterloo, ON N2L6P4 Canada, *p4li@math.uwaterloo.ca*; Jiahua Chen, University of British Columbia

Key Words: Asymptotic distribution, Exponential mixture, Finite mixture model, Likelihood ratio test, Score test

Testing for homogeneity in finite mixture models has attracted substantial research recently. Modified likelihood ratio test (MLRT) is a nice method because it has an asymptotically distribution-free test statistic and is locally most powerful. Interestingly, the mixture of exponential distributions or mixture models in scale distribution families do not satisfy the regularity conditions prescribed by many methods including the MLRT. To overcome this difficulty, we propose an iterative modified likelihood ratio test (IMLRT) in this paper. The IMLRT statistic has the same simple limiting distribution as MLRT statistic. The result is applicable to much more general mixture models and it does not require the parameter space to be bounded. Simulations show that the IMLRT has more accurate type I errors and higher powers under various models compared to existing methods.

The Equivalence Between Selecting and Hypothesis Testing

❖ Weixing Cai, Syracuse University, Department of Mathematics, 215 Carnegie Hall, Syracuse, NY 13210, *wcai01@syr.edu*; Pinyuen Chen, Syracuse University

Key Words: Correct selection, hypothesis testing, stepwise procedure, least favorable configuration,, Mahalanobis distance, operating characteristic functions

We consider various definitions of positive and negative populations, and define the corresponding "correct selection" (CS) for different selection goals. Depending on whether the covariance matrices are known or unknown, we propose single-step and stepwise procedures to achieve the goals of selecting populations that are equivalent to a given standard or a control by using Mahalanobis distance function for multivariate case(p>1). The univariate case will be considered as a special case where a standardized location distance is considered. The equivalence between selecting and hypothesis testing, including multiple hypothesis testing, is also established. Least favorable configurations are proved. Tables and graphs are presented to illustrate the properties of operating characteristic functions. Simulation examples will be given. Other forms of hypothesis testing will also be mentioned.

Small Sample Behavior of Resampling Methods

❖ Dongmei Li, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *dmli@stat.ohio-state.edu*; Jason C. Hsu, The Ohio State University

Key Words: Small sample behavior, Re-sampling techniques, Multiple testing

Re-sampling techniques are popular for analyzing microarray data. In this talk, we discuss the small sample behavior of three popular resampling techniques for multiple testing: permutation, post-pivot re-sampling, and pre-pivot resampling. We will show that when the sample size is small, permutation testing is unlikely to give small P-values; while post-pivot and pre-pivot resampling both might give P-values of zero for the same data, even adjusting for multiplicity. Conditions for the latter phenomenon will be described.

The Skew Generalized Secant Hyperbolic Distribution and Its Properties

❖ David Vaughan, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON N2L 3C5 Canada, *dvaughan@wlu.ca*

Key Words: generalized secant hyperbolic, skewness, model selection

Empirical studies, especially in finance, have shown that skew error distributions (typically with non-normal kurtosis) are common. The Generalized Secant Hyperbolic (GSH) is a symmetric class of distributions that has all moments finite (in contrast to Student's t), and with kurtosis ranging across all regular unimodal distributions. In this paper, skewness is introduced to the GSH in a natural way. Fundamental properties of the Skew Generalized Secant Hyperbolic (SGSH) and some basic model selection methods are discussed. This approach is contrasted with other methods for introducing skewness to symmetric distributions.

Generalized Exponentiated Beta Distribution

❖Alfred A. Akinsete, Marshall University, 1 John Marshall Drive, Department of Mathematics, Huntington, WV 25755, *akinsete@marshall.edu*

Key Words: Beta distribution, Distribution function, Entropy, Exponentiation, Moments, Parameter estimation

This paper studies a three-parameter generalization of the exponentiated beta distribution (GEBD). A special case of the distribution is found in Nadarajah (2005). We discuss various statistical properties of this distribution, and obtain expressions for the moment generating function, characteristic function, higher order moments; mean, variance, skewness and kurtosis. The hazard rate function of the distribution is discussed, and its asymptotic behavior is investigated. We demonstrate the sensitivity of the distribution to events where high risks are involved. The method of maximum likelihood estimation for the parameters of the distribution is discussed. The densities of the generalized exponentiation of other variants of the beta distribution are provided for future investigation.

A Study of Robust Estimation Approach for Analysis of Variance When the Distribution of Error Terms Is Non-Normal

❖ Aysun Çetinyürek, Ba¸kent University, Baglica Kampusu, ›statistik ve Bilgisayar Bilimleri Bolum, Ankara, 06530 Turkey, *cetinyurek@yahoo.com*; Bridal Senoglue, Ankara Universities

Key Words: modified likelihood, nonnormality, beta distribution, robustness

Analysis of Variance (ANOVA) is used to test the equality of the means when there are several groups. In ANOVA procedures, traditionally the error terms are assumed to be normally distributed. However, non-normal distributions are more prevalent in practice. In this paper, we derive the estimators of the model parameters in one-way and two-way ANOVA when the distribution of error terms is Beta. For Beta(a,b) distribution, the maximum likelihood (ML) method does not provide explicit estimators for the model parameters. Explicit estimators are derived via modified maximum likelihood (MML) methodology by linearizing the intractable terms in likelihood equations. MML estimators are known to be asymptotically fully efficient in terms of the minimum variance bounds (MVBs) and they are also robust. Hence, we propose to use MML estimators in ANOVA models when the error distribution is Beta(a,b).

27 Multiple Hypothesis Testing and Geomics ✪

IMS, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

Hierarchical Dependence in Meta-Analysis

❖ John R. Stevens, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, *john.r.stevens@usu.edu*; Alan M. Taylor, Brigham Young University-Idaho

Key Words: meta-analysis, dependence structure

A meta-analysis traditionally assumes the independence of study results. In cases where this assumption is violated, the dependence among study results can be estimated and accounted for at the sampling level in the meta-analysis to avoid overlapping information. We present an approach to account for the dependence at the hierarchical level also, effectively down-weighting extreme effect size estimates that are hierarchically dependent. This hierarchical dependence is estimated using both random effects and Bayesian models. We apply this approach to a meta-analysis of studies examining the effect of native language vocabulary aids on second language reading comprehension. Application to gene expression studies is also discussed.

Sequential Hypothesis Testing in Sensor Networks

❖ Yajun Mei, Georgia Institute of Technology, ISyE Georgia Tech, 765 Ferst Drive NW, Atlanta, GA 30332-0205, *ymei@isye.gatech.edu*

Key Words: distributed detection, sensor networks, sequential detection, Asymptotic optimality, quantization, tandem quantizer

The sequential hypothesis testing problem is studied in sensor networks, where a set of sensor receive independent observations and send summary messages to the fusion center, which makes a final decision. In the scenario where the sensors have full access to their past observations, the first of asymptotically optimal sequential tests in the literature is developed, and the proposed test has same asymptotic performance as the optimal centralized test that has access to all the sensor observations. Next, in the scenario where the sensors do not have full access to their past observations, a simple but asymptotically optimal sequential tests is developed, in which sensor message functions are what we call tandem quantizer, where each sensor only uses two different sensor quantizers with at most one switch between these two quantizers.

Modified Holm's Critical Values Using Correlation

❖ Zijiang Yang, Temple University, 2108B Mather Way, Elkins Park, PA 19027, *zjyang@temple.edu*; Sanat Sarkar, Temple University

Key Words: Multiple comparisons, Familywise error rate, Step-down test, Modified Holm's procedures

Holm (1979) proposed a step-down procedure for multiple hypothesis testing with a control of the familywise error rate (FWER) under any kind of dependence. Seneta and Chen (2005) sharpened the Holm procedure by taking into account the correlations between the test statistics. In this article, the Seneta-Chen procedure is further modified yielding a more powerful FWER controlling procedure. Simulations were carried out to show that our procedure often provides much better FWER control and power than the Seneta-Chen procedure.

On the Probability of Correct Selection for Large k Populations with Application to Microarray Data

❖ Jason Wilson, University of California, Riverside, Dept of Statistics, Riverside, CA 92521, *jason.wilson@email.ucr.edu*; Xinping Cui, University of California, Riverside

Key Words: Probability of Correct Selection, Ranking and Selection, Multiple Comparison, Microarray, d-best, G-best

One frontier of modern statistical research is the "multiple comparison problem" (MCP) arising from data sets with large k (>1000) populations (e.g., microarrays and neuroimaging data). In this talk we demonstrate an alternative to hypothesis testing. It is an extension of the Probability of Correct Selection (PCS) concept, which avoids the MCP by its nature. The idea is to select the top t out of k populations and estimate the probability that the selection is correct, according to specified selection criteria. We propose "d-best" and "G-best" selection criteria that are suitable for large k problems and illustrate the application of the proposed method on two microarray data sets. Results show that our method is a powerful method for the purpose of selecting the "top t best" out of k populations.

Alpha-Investing: A New Multiple Hypothesis Testing Procedure

❖ Robert Stine, University of Pennsylvania, 444 Huntsman Hall, Department of Statistics, Philadelphia, PA 19104-6340, *stine@wharton.upenn.edu*; Dean Foster, University of Pennsylvania

Key Words: Bonferroni, False discovery rate, FDR, Multiple comparisons, family-wide error rate

Alpha-investing is an adaptive, sequential methodology that encompasses a large family of procedures. All control mFDR, the ratio of the expected number of false rejections to the expected number of rejections. mFDR is a weaker criterion than FDR, which is the expected value of the ratio. We compensate for this weakness by showing that alpha-investing controls mFDR at every rejected hypothesis. Alpha-investing resembles alpha-spending used in sequential trials, but possess a key difference. When a test rejects a null hypothesis, alpha-investing earns additional probability toward subsequent tests. Alpha-investing hence allows one to incorporate domain knowledge into the testing procedure and improve the power of the tests. In this way, alpha-investing enables the statistician to design a testing procedure for a specific problem while guaranteeing control of mFDR.

Statistics on High-Dimensional Data Analysis

❖ Yingli Qin, Iowa State University, 315C Snedecor Hall, Ames, IA 50010, *qinyl@iastate.edu*; Song Xi Chen, Iowa State University

Key Words: high dimensional data, FDR, multiple testing, Bonferroni

This talk is dealing with high-dimensional data analysis. I will start with two sample test. Theoretical results will be given briefly because they have been reported in my previous talk on JSM 2006. This time I will mainly focus on the simulation analysis. The simulation will be designed to compare BS test and our new test with FDR procedure and Bonferroni procedure. I will proceed to introduce more test problems on high-dimensional data analysis if possible.

Pseudo-ROC: Method Comparisons in the Absence of a 'Gold Standard'

❖Richard Bourgon, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, CB5 8HP United Kingdom, *bourgon@ebi.ac.uk*

Key Words: Sensitivity/specificity, Receiver operating characteristic curve, Gold standard, Diagnostic test

In many applications in quantitative biology, one seeks to compare different statistical or laboratory procedures on the basis of sensitivity and specificity. Frequently, large "gold standard" test sets are not available, but contaminated test sets—in which putative true positives contain some true negatives, and vice versa—are plentiful. Various authors have proposed methods for estimation of sensitivity and specificity, or alternatively, relative true and false positive rates, in this setting. Most assume conditional independence of procedures given the true status of each test set member, but this is often unrealistic. Here, we describe a novel "pseudo-ROC" approach which is intuitively appealing and is valid under weaker independence assumptions. Examples from ChIP-chip microarray experiments are provided, along with a simple statistical test for equivalence of pseudo-ROC curves.

28 Health Policy Interventions and Methods ● ✪

Section on Health Policy Statistics, WNAR

Sunday, July 29, 2:00 pm–3:50 pm

Relationship Between Competitiveness of Colleges and Levels of Tobacco Use

❖ Joost Bottenbley, 7001 Birkengead Pl, Alexandria, VA 22315, *yoshiz008@yahoo.com*; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University; Samer Ellahham, Innovative Medical Institute/Paragon Cardiovascular Foundation; David Anderson,

Key Words: Tobacco, Competitiveness, Colleges

This paper describes the relationship between the compositeness of a school and the level of smoking and makes recommendations on which category of smokers anti-smoking campaigns can target to reduce the level of smoking the most for the least amount of money. Data was collected from 15 schools over a period of three years (f2003 to 2005) in the state of Colorado.

Restriction of Trans Fatty Acids in Washington, DC, Metropolitan Area: Health Benefits and Economic Impact

Joost Joost Bottenblehy, Johns Hopkins University; Samer Ellahham, Innovative Medical Institute/Paragon Cardiovascular Foundation; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University; ❖Malda Bashi, Paragon Cardiovascular Foundation, , *samer1@ix.netcom.com*

Key Words: partially hydrogenated oils, economic effects, heart disease

Heart disease is the leading cause of death in the United States. Scientific evidence demonstrates a clear association between increased trans fat intake and the risk of coronary heart disease. Transfatty acids are geometrical isomers of the naturally occurring cis-fatty acids. Recently the City of New York has restricted the use of all trans fatty acids in foods. The economic impact of such a restriction in the Washington, DC area is evaluated. Five possible options for restaurant owners and corporations to reduce the negative health risks of trans fatty acids are evaluated. The model includes health benefits, consumption rate, geographic and economic impact.

Price Indexes with a Health Insurance Component

❖Ralph Bradley, Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Room 3105, Washington, DC 20212, *bradley.ralph@bls.gov*

Key Words: risk aversion, price index, imputation

Currently, the Consumer Price Index (CPI) does not have a sub index for health insurance. Instead, the effects of health insurance are imputed using the prices of the medical goods and services that are in the CPI sample. This imputation method assumes that changes in health insurance premiums can be well predicted by changes in the prices of the medical items that health insurance covers. Unfortunately, there is much evidence that the growth of health insurance premiums have consistently outpaced the growth in both medical prices and utilization. If that evidence is correct, the CPI is not correctly incorporating the effects of health insurance premium growth. In this study, I price the income protection services of the health insurance industry and I incorporate this into an alternative CPI and compare it to the published CPI.

Multiple Imputation of the Euroqol EQ-5D: Testing Different Approaches Using IVEWare

❖Robert Gerzoff, Centers for Disease Control and Prevention, MS K10, Atlanta, GA 30341, *rcg8@cdc.gov*; Theodore J. Thompson, Centers for Disease Control and Prevention

Key Words: Imputation, Diabetes, Managed Care

The EQ-5D is a standardized and widely used health index consisting of five items each with three levels. Different response profiles are converted to a single index value using established valuation weights. When individual items are missing the analyst must decide whether to impute the EQ-5D directly, or impute the missing item(s) and recalculate the EQ-5D. Decisions must also be made about the distributional assumptions underlying the missing data and the alternative methods available to model the missingness. Using data from the TRIAD Study, a multicenter study of diabetic 11,927 patients enrolled in ten U.S. managed care plans in seven states, we simulate missing data patterns and use IVEW are to evaluate the performance of various imputation strategies.

A Choice of Prediction Rules in Logistic Regression Models

❖Melvin Ott, 2613 E Nicklaus Ave, Spokane, WA 99223, *ott@asisna.com*

Key Words: logistic regression, prediction rules, health care, market research, banking

Several authors have suggested prediction rules for use in logistic regression models. However, no agreed standard for selecting prediction rules from a logistic regression model has been given. Neter et al., in his text Applied Linear Regression Models, suggests three rules: 1. “Use .5 as the cutoff.” 2. “Find the best cutoff for the data set on which the multiple logistic regression model is based.” 3. Use prior probabilities and costs of incorrect predictions in determining the cutoff. Halpern, Gallop, and others have suggested finding an optimal operating point (OOP) from the ROC curve. The discussion from this paper suggests that there is no “gold” standard for selecting the prediction rule for all applications. Three datasets and logistic models will be presented representing market research, health care, and banking applications.

Nonparametric Confidence Intervals for the Mean of a Right-Skewed Distribution

❖Kenny Shum, Johns Hopkins University, 2100 Lee Highway Apt123, Arlington, VA 22201, *kshum@jhsph.edu*; Scott Zeger, Johns Hopkins University

Key Words: Skewed distribution, Medical cost, Bootstrap, Confidence interval, Mean estimation

The skewness in the distribution affects both the performance of the sample mean and the coverage of standard confidence intervals in small to medium sample size. Transformation and bootstrap methods have been proposed to improve coverage of the approximate intervals. We compare the finite sample performance of confidence intervals for the mean of a class of right skewed distribution ranging gamma to lognormal. We highlight the improvement with the use of bootstrap calibration. This work is motivated by the estimation of the mean medical cost.

29 Data Analysis and Modeling ● ✪

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Sunday, July 29, 2:00 pm–3:50 pm

A New Statistical Approach for the Analysis of Uncertain Systems

❖Xinjia Chen, Louisiana State University, Dept of ECE, Baton Rouge, LA 70803, *chan@ece.lsu.edu*; Kemin Zhou, Louisiana State University; Jorge Aravena, Louisiana State University

Key Words: Monte Carlo Simulation, Sample Reuse, Robust Statistics, Numerical Method, Uncertainty, Control Systems

This paper addresses the issues of conservativeness and computational complexity of probabilistic robustness analysis. We solve both issues by defining a new sampling strategy and robustness measure. The new measure is shown to be much less conservative than the existing one. The new sampling strategy enables the definition of efficient hierarchical sample reuse algorithms that reduce significantly the computational complexity and make it independent of the dimension of the uncertainty space. Moreover, we show that there exists a one to one correspondence between the new and the existing robustness measures and provide a computationally simple algorithm to derive one from the other.

The Efficiency of Ridge Regression Estimators with Respect to Zellner's Balanced Loss Function

❖ Marvin Gruber, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623, *mjgsma@rit.edu*

Key Words: balanced loss function, matrix loss, simultaneous estimation, total mean square error, ridge estimator

The optimum ridge estimator is obtained for a setup with r linear models for a generalization of Zellner's balanced loss function. Two important special cases are considered, matrix loss and the total mean square error. For each of these two cases an estimator of the biasing parameter matrix is obtained that produces a mean square error that is uniformly smaller than that of the least square estimator. The results are illustrated with an example that uses numerical data. In general the best improvement in the MSE over the least square estimator occurs for the classical loss function. Parameters with smaller values produce smaller MSE.

Identification of Dispersion Effects in Replicated Experiments

❖ Cheryl Dingus, Battelle, 505 King Ave, 11 7 026, Columbus, OH 43201, *dingusc@battelle.org*; Angela Dean, The Ohio State University; Bruce Ankenman, Northwestern University

Key Words: Dispersion, Replicated Experiment

A factor is said to have a dispersion effect if the variability of the response is different at the different factor level settings. In replicated experiments, replicate observations for a treatment are used as a measure of the variability of response. Using traditional methods, a summary statistic is used as the analysis variable in an F-test. The disadvantage of traditional analysis methods is the loss of power to identify effects. This talk will present an alternative method of analysis that transforms each observation into an "individual measure of the variability", preserving all the original degrees of freedom and thereby increasing power over the traditional method. Several functions that have been examined for this alternative analysis will be presented. Results will show that the new methodology is better able to detect dispersion of a single factor than the traditional method.

A Sequential Methodology for Change Point Estimation

❖ Yan Lan, University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109, *lany@umich.edu*; Moulinath Banerjee, University of Michigan; George Michailidis, The University of Michigan

Key Words: adaptive sampling, change point estimation, two-stage procedure, Skorohod Topology

Consider a constant regression model for a bounded covariate that has a single discontinuity. Budget constraints dictate that a total of N covariate values and corresponding responses can be obtained. The goal is to estimate accurately the location of the change-point. We propose a two-stage procedure and its properties are examined, where at the first stage a proportion of the N points are sampled and the location of the change-point estimated. Subsequently, the remaining proportion of points are sampled from an appropriately chosen neighborhood of the initial estimate and a new estimate is obtained. The asymptotic statistic of the least squares estimate is derived. The rate of convergence is improved to o(1/n^{1+\gamma}) (\gamma\in(0, 1)).The improved efficiency of the procedure is demonstrated using real and synthetic data. The methodology could be generalized to a multi-stage procedure.

Inference for Multipiecewise Regression Using the Bootstrap

❖ Trevor Craney, Sikorsky, 13 Belinsky Circle, Oxford, CT 06478, *trevor.craney@sbcglobal.net*

Key Words: changepoint, pants

Inference in ordinary least squares regression is dependent on various assumptions, including normality. When the least squares approach is constrained, as it is in piecewise regression, the standard assumptions can no longer be assumed valid. The bootstrap approach is applicable to creating confidence, or inference, in this case. Further, if the user allows more than one changepoint to be estimated by the data, a suitable stopping method must be decided, which is not unlike the question of how many variables to add in a multiple regression model using a best subsets or forward selection approach. The bootstrap is shown to still be applicable in the multi-piecewise regression model with various bootstrap approaches to be considered by the analyst.

Generalized Confidence Intervals for Proportion of Conformance

❖ Paul Patterson, USDA Forest Service, 2150A Centre Ave, Suite 350, Fort Collins, CO 80526, *plpatterson@fs.fed.us*

Key Words: Proportion of conformance, Confidence intervals, Fiducial generalized pivotal quantity, Quality control

For a random variable X, with c.d.f. F, the proportion of conformance for an interval (A,B) is F(B)-F(A). There are many practical applications, such as quality control and environmental monitoring, where confidence intervals for the proportion of conformance are needed. In this paper a method is presented for constructing approximate confidence intervals for the proportion of conformance for two distributions: the normal distribution with unknown mean and variance and the uniform distribution with unknown location and scale parameters. The method is very general and can easily be extended to other distributions. The approximate confidence intervals are constructed using Fiducial Generalized Pivotal Quantities. Performance of the one-sided limits is assessed using a simulation study. The asymptotic exactness of the confidence interval coverage will be discussed.

Analysis of Probability of Detection and Probability of Alarm on Sonic IR Inspection Data Using the Mixed Effects Model and Bootstrapping

❖ Chunwang Gao, Iowa State University, 46 Schilletter VLG APT D, Ames, IA 50010, *cgao@iastate.edu*; William Meeker, Iowa State University; Donna Mayton, SAIC

Key Words: Sonic IR, Thermal Acoustic, Mixed Effect Model, Bootstrapping, Nondestructive Evaluation

Sequences of infrared images were taken by Sonic IR system on 10 cracks in 7 different samples. For each crack, different levels of vibration amplitude (3 levels), pulse length (3 levels) and trigger force (3 levels) of the Sonic IR system are arranged to study the relationship between the maximum contrast of the images and the experimental setup. Time to alarm was also recorded for those having an alarm. A quadratic random effects response surface model to the combined contrast data from the 10 cracks was fitted to provide an assessment of probability of detection as a function of the experimental variables. Another linear response surface model to the combined time-to-alarm data from the 10 blade cracks are fitted to provide an assessment of the probability of alarm. Bootstrapping is used to explore methods of quantifying statistical uncertainty in the two models.

30 Statistical Process Control and Other Related Topics in Quality ●✪

Section on Quality and Productivity

Sunday, July 29, 2:00 pm–3:50 pm

Multivariate Process Control for Improved Detection of Changes in Process Location

❖ Amitava Mitra, Auburn University, College of Business, Auburn, AL 36849-5240, *mitraam@auburn.edu*

Key Words: Process control, Cause identification, Average run length, Type II error

Much of multivariate process control has focused on monitoring product/process output characteristics. However, the output variables from a process are influenced by the levels of the input or regulatory variables and uncontrollable factors. Thus, traditional approaches may not identify which of the causal factors, i.e., process regulatory variables or uncontrollable noise factors or both, are possible reasons for out-of-control signals. Further, when the effect of the process regulatory variables and the noise factors cancel each other, the impact on the output response variables may not be detectable. Here, we consider the impact of both process regulatory variables and noise factors through an adjusted measure. The performance of the proposed method is determined, through simulation, by the average run length of first detection of out-of-control conditions.

Distribution-Free Filters Used To Determine Phase I Control Chart Limits for Spread and Location

❖ Victoria Jordan, Auburn University, 216 Parker Hall, Auburn, AL 36849, *jordavs@auburn.edu*; Allison Jones-Farmer, Auburn University; Charles Champ, Georgia Southern University

Key Words: Control Limits, Phase I, Distribution-Free Statistical Quality Control

SQC depends on the proper establishment of control limits in Phase I. Current Phase I procedures are based on "trial and error" and many require normally distributed observations. Some robust and distribution-free control charts have been developed, but most are concerned with the establishment of Phase II limits, robust to non-normality or outliers in the Phase I sample. This research proposes a distribution-free method for filtering out data points associated with assignable causes in Phase I. The purpose is to identify an in-control reference sample that can be used to estimate the process parameters for spread and location for Phase II. The spread chart is based on a variation of Levene's test and the location chart is based on the Kruskal-Wallis signed-rank procedure. Monte Carlo simulation is used to compare performance with existing normal theory and distribution-free procedures.

Monitoring the Slopes of Linear Profiles

❖ Junjia Zhu, The Pennsylvania State University, 331A Thomas Building, University Park, PA 16802, *junjia@psu.edu*

Key Words: Profile Monitoring, Simple Linear Regression, Statistical Process Control

In this paper, we focus on monitoring the slopes of linear profiles. A Shewhart-type control chart for monitoring slopes of linear profiles is proposed. Both Phase I and Phase II applications are discussed. The performance of the proposed chart in Phase I applications is demonstrated using both real-life data in an illustrative example and simulated data in a probability of signal study. For Phase II applications, the average run length (ARL) is used for comparison. It is shown that the ARL of the proposed control chart depends only on the shifts of slopes; whereas the ARL of the multivariate T2 chart depends on both the shifts of slopes and the correlation between the estimated slope and the intercept. When such a correlation is low (say within +/-0.6), the proposed control chart has a better ARL performance than the T2 chart.

Monitoring of Hazard Rates for Dynamically Changing Observations

❖ Emmanuel Yashchin, IBM T.J. Watson Research Center, Box 218, Yorktown Heights, NY 10598, *yashchi@us.ibm.com*

Key Words: Warranty, Reliability, Control Charts, Cusum, Change Point, Wearout

We consider life testing situations in which a relatively large portion of data that serves as a basis for a monitoring scheme changes with time. Situations of this type are common in the areas of accelerated life testing or warranty data analysis. One of the key problems in such situations is detection of changes in the hazard rate for the tested population. For example, in the warranty data setting one may want to detect, as quickly as possible, onset of excessive wearout conditions for some set of manufacturing vintages. Since the whole data set is affected at each point in time that new information is obtained, one needs a special approach to the problem of monitoring lifetime characteristics. In this paper we discuss one such approach and give an example related to the problem of detection of trends in warranty data for electronic components.

Benefits of Integrating Operational Profiles with Classical Acceptance Sampling in the 2010 U.S. Census

❖ Thomas George, Evolver, Inc, 7833 Walker Drive, Greenbelt, MD 20770, *tgeorge1@evolverinc.com*

Key Words: Bayesian, Quality, Production, Keying, OCR

Although classical acceptance sampling methodologies ensure a specified output quality level, the cost of achieving that output quality is a function of the input quality process, which is specific to each application. However, classical sampling methodology forecasts are general in scope. Further, some assessment of input quality is necessary in order to generate a system that will produce an output quality with verifiable costs. Once established, if the input to a system changes, frequently the only way to meet output quality levels are changes in operations. Therefore, the methods used for integrating operational and acceptance sampling profiles outlined in this paper allowed for improved cost forecasting in regards to quality at all stages of a system's life for the U.S. 2010 Census.

Sampling Procedures for Extending the Use of the Fishbone Diagram

❖ Frank Matejcik, South Dakota School of Mines, Industrial Engineering, 501 E Saint Joseph Street, Rapid City, SD 57701-3901, *frank.matejcik@sdsmt.edu*

Key Words: Ishikawa, Fishbone, Capture, Recapture

Procedures for creating a Fishbone diagram often assume all the possible participants are available. However, sampling among the participants may be required. The sampling problem is similar to a capture recapture. Methods for sampling decisions and experiences from classroom trials are discussed. Potential applications are mentioned, also.

31 Interesting Statistical Applications

Section on Statistical Education, General Methodology, Section on Teaching Statistics in the Health Sciences

Sunday, July 29, 2:00 pm–3:50 pm

Regression Through the Origin: Models and Methods

❖ Patrick Johnston, Children's Hospital Boston, Clinical Reseach Program, 300 Longwood Ave, Boston, MA 02115, *patrick.johnston@childrens.harvard.edu*

Key Words: regression through the origin, generalized linear models, imputation

Bowel lengths for infants undergoing gastrointestinal surgery are commonly imputed by more readily available lengths for deceased infants of a similar gestational age. For reasons of tissue elasticity, it is reasonable to expect that the latter overestimate the former by a constant factor. Several simple and intuitive methods have been proposed to estimate this proportionality factor, and we relate these estimates to more general model-based estimates. Models we consider are regressions through the origin estimated by maximum likelihood, maximum quasi-likelihood, and least squares. Aspects related to efficiency, robustness, and parameterization are discussed.

Testing the Score Invariance of State Performance Tests Over Time

❖ Jiahe Qian, Educational Testing Service, Rosedale Road 02T, Princeton, NJ 08541, *jqian@ets.org*

Key Words: test score inflation, proficiency standards, assessment accountability, mapping, NAEP, NCLB

In recent years, test score inflation has compromised efforts to improve education and the accountability of performance assessments. One assessment with potential for test score inflation is state performance tests. Under the No Child Left Behind Act (NCLB), all states are required to select the tests and set the proficiency standards in reading and mathematics for grades 3-8 as well as in at least one grade in high school. This study develops a statistical approach to examine score inflation in state tests. First, the state proficiency standards are mapped to the National Assessment of Educational Progress (NAEP) scale. Then, rather than testing inflation of scores, we test the invariance of the mapped proficiency standards of state tests over time. The mapping technique is based on the enhanced method developed by Braun and Qian (2006).

Inference for Normal Mixture Model, Univariate or Multivariate, in Both Mean, and Variance

❖ Xianming Tan, Queens University, 10 Stuart Street, NCIC Clinical Trials Group, Kingston, ON K7L3N6 Canada, *xtan@ctg.queensu.ca*; Jiahua Chen, University of British Columbia

Key Words: Bernstein inequality, invariant estimation, mixture of normal distributions, penalized maximum likelihood, strong consistency

Due to the unboundedness of likelihood function, the statistical inference based on data from the finite mixture of normal distributions in both mean and variance is particularly difficult. Through the use of some suitable penalized likelihood functions, we develop a new method applicable to both univariate and multivariate normal mixture models. We show that the penalized maximum likelihood estimators (PMLE) are strongly consistent, asymptotically normal, and invariant under linear transformations. We also show that the constrained maximum likelihood estimator, as proposed in Hathaway (1985), is still consistent when the lower bound c placed on the ratio of any two component variances satisfies that $c \geq \exp \{-k (\log^2n)\}, ~ \forall ~ k>0$. Furthermore, this approach can be readily extended to solve similar problems in other mixture models.

Measurement Error in Factor Analysis: The Question of Structural Validity

❖ Michael K. Lauritzen, Brigham Young University, 1122 SWKT, Provo, UT 84602, *mikelauritzen@byu.edu*; Landon Poppleton, Brigham Young University; Mitchell Harris, Brigham Young University; Naomi Hunsaker, Brigham Young University; Robert Bubb, Brigham Young University; Bruce Brown, Brigham Young University

Key Words: Factor analysis, Structural integrity, Reliability, Monte Carlo simulation

The central claim of factor analysis is that it can identify underlying source variables from the correlational pattern in a larger set of surface variables. Monte Carlo simulation is used to create surface variables from a systematic underlying structure, either two-dimensional or three-dimensional, either clustered or Toeplitz, with one of six levels of measurement error added to the data. Quantitative and also graphical methods are used to assess the extent to which structural integrity is maintained for each type of data at each level of measurement error. The watershed value is a reliability of .50. Below that level the structural integrity of the solution deteriorates rapidly. Interestingly, in a survey of papers in two top psychology journals that used factor analysis, not one of the forty-three studies in the sample had adequate reliability levels to provide structural integrity.

The Central Limit Theorem and Structural Validity in Factor Analysis

❖ Landon Poppleton, Brigham Young University, 1001 SWKT, Provo, UT 84602, *landonpoppleton@hotmail.com*; Mitchell Harris, Brigham Young University; Michael K. Lauritzen, Brigham Young University; Robert Bubb, Brigham Young University; Bruce Brown, Brigham Young University

Key Words: factor analysis, optimization, central limit theorem

Factoring methods such as principal components and factor analysis are often used for data simplification. Over the past century a number of strategies have evolved for optimizing factor analysis methods. The received view consists of recommendations for various extraction or rotational methods, recommended sample size, etc. An alternative view is proposed. It is argued that the single most important consideration in producing structurally valid factor solutions is the reliability of individual variables, and that an approach based upon the central limit theorem is particularly effective in increasing data reliability. Using a Monte Carlo simulation approach it is demonstrated that the central limit theorem averaging method is substantially more effective in producing structurally valid analyses than the various recommendations of the received view.

Determining a Minimum Number of Observations in Factor Analytic Studies

❖ Robert Bubb, Brigham Young University, 47 Wymount, Provo, UT 84604, *robert_bubb@byu.edu*; Bruce Brown, Brigham Young University

Key Words: Factor Analysis, Monte Carlo, Variable Stability

The minimum number of observations necessary in factor analytic studies to produce latent variable stability is inconsistent and highly debatable. Research studies recommend minimum observations ranging from 100 to 1000 observations or observation-to-variable ratios of 2:1 to 20:1. Research has also demonstrated that the number of observations is clearly not sufficient in itself to produce latent variable stability. A Monte Carlo simulation method is used to manipulate the number of observations, number of factors, ratios of errorless-to-random data, and the number of variables-to-factors. MANOVA statistics are used to determine latent variable stability given the various conditions. Results indicate an intricate relationship between observations, factors, degree of error, and variables. Under certain conditions, 16 observations are sufficient to produce latent variable stability.

A Meta-Analysis of the Content and Effectiveness of Training Programs for Federal and State Transportation Employees

John Aleong, University of Vermont; ❖Chandra Aleong, Delaware State University, 1200 DuPont Highway, Dover, DE 19901, *caleong@desu.edu*

Key Words: Training programs, benchmarking, evaluation, extrapolation, Transportation, Federal and State

Successful tracking of training programs and their effectiveness is necessary to determine whether transportation professionals are being prepared for the challenges of the 21st century. Training is integral for the execution of strategies at the federal and state levels and millions of dollars are spent on various forms of educational opportunities. The issues that affect transportation are very diverse and span disciplines and international boundaries. A major research question is "Are taxpayers getting their money's worth and are employees getting the necessary skills for the new environment?" This topic is also important for study given the high rate of retirement and downsizing taking place across the country in all sectors of the economy. Outsourcing has become commonplace and is resulting in the loss of significant intellectual capital.

32 Water-Related Issues and Methodology ●

Section on Statistics and the Environment, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

Functional Linear Models for Daily and Yearly Streamflow Measures

❖Mark Greenwood, Montana State University, Department of Math Sciences, PO BOX 172400, Bozeman, MT 59717-2400, *greenwood@math.montana.edu*; Johnnie N. Moore, University of Montana; Joel T. Harper, University of Montana

Key Words: functional data analysis, functional linear model, hydrology

Recent analyses of regional snow and river data for the late 20th century suggests that western North American snow pack has decreased (Mote et al. 2003) and the spring snow melt pulse has arrived earlier (Stewart et al. 2005). Stewart et al. (2005) found significant trends in various measures of spring runoff timing. This work analyzes streamflow runoff using functional linear models (Ramsay & Silverman, 2005) using data from 1951 to 2005 from 21 gages located on the Missouri and Columbia headwaters. Performing the analysis based on yearly summary measures is compared to analyzing daily flow measurements over the same time period. Potential inferences to the impacts of some forcing functions such as PDO and ENSO will also be explored.

Modeling Shark Bycatch: The Zero-Inflated Negative Binomial Regression Model with Smoothing

❖Mihoko Minami, Institute of Statistical Mathematics, 4 6 7 Minami Azabu Minato ku, Tokyo, 106 8569 Japan, *mminami@ism.ac.jp*

Key Words: GAM, GLIM, thin plate regression splines, Generalized Information Criteria, EM algorithm, Partial dependence

The zero-inflated negative binomial (ZINB) regression model with smoothing is introduced for modeling count data with many zero-valued observations, and its use is illustrated with shark bycatch data from the eastern Pacific Ocean tuna purse-seine fishery. Based on the generalized information criterion, the ZINB regression model provided a better fit to the data than either Poisson, negative binomial or ZIP regression models. To demonstrate the utility of the ZINB regression model for the standardization of catch data, standardized temporal trends in bycatch rates estimated with the ZINB regression model are computed and compared to those obtained from fits of the other three types of models. Comparison of trends among models suggests that the negative binomial regression model may overestimate model coefficients when fitted to data with many zero-valued observations.

A Distance-Based Classifier with Application to Microbial Source Tracking

❖Jayson Wilbur, Worcester Polytechnic Institute, Dept of Mathematical Sciences, 100 Institute Road, Worcester, MA 01609, *jwilbur@wpi.edu*

Most classification rules can be expressed in terms of distances from the point to be classified to each of the candidate classes. For example, linear discriminant analysis classifies points into the class for which the sample Mahalanobis distance is smallest. However, dependence among these point-to-group distance measures is generally ignored. In this talk, a general classification rule will be defined which uses information about this dependence structure to improve classification. This work was initially motivated by the problem of microbial source tracking which aims to identify sources of fecal contamination in water resources based on genotypic and phenotypic variation in public health indicator organisms such as E. coli. An application of the proposed methodology to microbial source tracking will be presented.

Grain Size Distributions in Modern and Ancient Fluvial Systems

❖Snehalata Huzurbazar, University of Wyoming, Dept 3332, 1000 E. University Avenue, Laramie, WY 82071, *lata@uwyo.edu*

Key Words: hydrology, functional data analysis

Examining grain-size distributions of deposits in rivers is an important component in understanding sediment transport systems. We analyze data obtained from a high-resolution laser analysis of particle sizes in samples obtained from three modern rivers and three ancient rivers. Exploratory data analysis is carried out to examine possible clusters and variation in the empirical distributions of the grain sizes. We next discuss modeling the data using hyperbolic distributions.

Spatial Modeling for Groundwater Arsenic Levels in North Carolina

❖Dohyeong Kim, Duke University, Box 90328, CEHI, Durham, NC 27708-0328, *dohyeong@duke.edu*; Marie Lynn Miranda, Duke University; Joshua Tootoo, Duke University; Alan E. Gelfand, Duke University

Key Words: arsenic in groundwater, spatial random effects model, 3-dimensional arsenic surface, GIS

Growing evidence on the relationship between long-term exposure to arsenic in drinking water and negative health outcomes has drawn attention to unregulated private wells with potentially high levels of arsenic. Little is known about the environmental determinants of arsenic in groundwater. Data collected from 475 private wells in Orange County, North Carolina were integrated with detailed geological and tax assessor's parcel data into a GIS. The data were initially fitted to a left-censored regression model to identify key determinants of arsenic levels in groundwater. Then, a Bayesian spatial random effects model was developed to illuminate spatial pattern in arsenic residuals. The resulting models provide a basis for public health intervention by creating three-dimensional surface maps of predicted arsenic levels in groundwater for any location and depth in the area.

Change Point Analysis for Water Pressure Data

Snehalata Huzurbazar, University of Wyoming; ❖Arunendu Chatterjee, University of Wyoming, 354 Soule Street Apt 287, Laramie, WY 82072, *achatte1@uwyo.edu*

Key Words: glacial hydrology, change points, wavelet coefficients

Our objective is to identify and model change points in the measurements of sub-glacial water pressure during melt season along the length of the Bench glacier in Alaska. This modeling will provide insights into the subglacial hydrology including the discharge mechanism during the melt season. As a first step in this modeling, we consider the identification of change points in the twenty water pressure data sets available. We compare an existing wavelet based method suggested by Wang (1995), which examines the empirical wavelet coefficients of the data at the fine scale levels with a Bayesian method based on the empirical wavelet coefficients for the standard change point problem by Ogden (1996). The above methods have to be adapted for accommodating missing observations, which are present in our datasets.

Weibull Prediction Limits with Retesting

❖Kirk Cameron, MacStat Consulting, Ltd, 10330 Mill Creek Ct, Colorado Springs, CO 80908, *kcmacstat@qwest.net*

Key Words: Weibull distribution, prediction limits, retesting, groundwater, monitoring

EPA guidance for groundwater monitoring recommends simultaneous prediction limits combined with retesting. Under this rubric, an initial exceedance of the background prediction limit among measurements from a set of monitoring wells is followed by a confirmatory test at each exceeding well. This allows control of nominal site-wide false positive rates (SWFPPR) while ensuring power to detect contaminated groundwater. Past research has led to solutions for normal-, lognormal-, and gamma-based limits, as well as non-parametric limits. We extend that work to simultaneous two-parameter Weibull prediction limits, illustrating the theory and providing methods for practical computation. We also compare Weibull limits against other parametric and nonparametric limits in real samples of groundwater data. The Weibull distribution acquits itself as a flexible tool in groundwater monitoring.

33 Modeling and Estimation in Environmental and Occupational Health ●

Section on Statistics in Epidemiology, Section on Health Policy Statistics

Sunday, July 29, 2:00 pm–3:50 pm

A Comparative Analysis of the Chronic Effects of Fine Particulate Matter

❖Sorina Eftim, Johns Hopkins University, 607 9th Street, Laurel, MD 20707, *seftim@jhsph.edu*; Holly Janes, Johns Hopkins University; Aidan McDermott, Johns Hopkins University; Jonathan Samet, Johns Hopkins University; Francesca Dominici, Johns Hopkins University

Key Words: air pollution, fine particulate matter, mortality, Medicare, chronic effects, elderly

The American Cancer Society (ACS) study and the Harvard Six Cities study (SCS) are the two landmark cohort studies for estimating the chronic effects of fine particulate matter (PM2.5) on mortality. To date, no comparative analysis of these studies has been carried out using a different study design, study period, data, or modeling approach. We estimate the chronic effects of PM2.5 on mortality for the period 2000–2002 using mortality data from Medicare and PM2.5 levels from the National Air Pollution Monitoring Network for the same counties included in the SCS and the ACS. We use a log-linear regression model which controls for individual-level and area-level covariates. The results are similar to those reported by the original SCS and ACS indicating that fine particulate matter is still significantly associated with mortality when more recent air pollution and mortality data are used.

The Effect of PM2.5 on Hospital Admissions Rates by Season and Region

❖Keita Ebisu, Yale University, 60 College Street, New Haven, CT 06510, *zhize@aol.com*; Michelle Bell, Yale University; Roger Peng, Johns Hopkins University; Francesca Dominici, Johns Hopkins University

Key Words: Air Pollution, Hospital Admissions, Particulate Matter

We investigated the seasonal and regional associations between PM2.5 levels and cause-specific hospital admission rates using two models: a pollutant-seasonal indicator model; and a model to smooth seasonal patterns of county-specific log relative rates with a sine/cosine function. Both models were fitted using a generalized additive model adjusting for potential confounders. In order to pool the county-specific estimates to form national average estimates, we fit a Bayesian hierarchical model using Two-Level Normal independent sampling estimation with non-informative priors. The results of both models showed that the association between PM2.5 and admissions were larger in the winter for cardiovascular and respiratory causes. There was also a regional difference between eastern and western US.

Statistical Issues in Health Effects of Air Pollution

❖Donna Stroup, Klemm Analysis Group, Inc., 1339 Green Court, 4th Floor, Washington, DC 20005, *rklemm@klemmanalysis.com*; Rebecca Klemm, Klemm Analysis Group, Inc.

Key Words: air pollution, particulate matter, modeling, environmental epidemiology, distance effect

We examine issues in estimating the effect of air pollution on the health of populations. We used mortality data from CDC, air quality data from EPRI and EPA, and weather data from NOAA. We calculate distance between location of death and monitoring and the sensitivity of estimates to this distance and smoothing parameters. Multiple cause of death (in addition to underlying cause) adds additional information for selected causes of death, particularly respiratory conditions, unintentional injuries, septicemia, and Alzheimer's disease. Choices of parameters in modeling were less sensitive to distance from power monitoring source than previously determined. Proper use and interpretation of models of air pollution's effect on human health should enable more complete understanding of the effect of air pollution and should inform policy in this arena.

Exposure-Response Assessment of Airborne Indoor Allergen Immunoassay

❖Juan Liang, University of Cincinnati, Department of Environmental Health, PO Box 670056, Cincinnati, OH 45267, *liangji@email.uc.edu*; Linda Levin, University of Cincinnati; Tiina Reponen, University of Cincinnati; Grace LeMasters, University of Cincinnati; David Bernstein, University of Cincinnati; Rolanda Olds, University of Cincinnati; Zana Lummus, University of Cincinnati

Key Words: Four Parameter Logistic, Nonlinear Mixed Model, Immunoassay

This study is to apply state-of-the-art statistical techniques to the analysis of allergen assay data. Ultimately it will unify statistical methodology for estimating allergen concentrations. Data were from the Cincinnati Childhood Allergy and Air Pollution Study (CCAAPS). As part of a birth cohort study, dust samples from homes of infants in the Greater Cincinnati Area were analyzed by using monoclonal ELISA. Four-parameter logistic models were chosen. NLMIXED with heterogeneous variance structure was utilized to describe within- and between-assay variability in order to determine the standard exposure-response curves. By reversing the estimated standard curves, detection limits and calibrated concentrations of allergens were obtained. It was shown that the precision of estimated concentrations was substantially improved and the limits of the detection were quantified.

Correcting for Shared Measurement Error in Complex Dosimetry Systems

❖Terri K. Johnson, University of Southern California, 1540 Alcazar St CHP220, Los Angeles, CA 90089, *tkang@usc.edu*; Daniel Stram, University of Southern California

Key Words: Monte-Carlo maximum likelihood, full parametric bootstrap, complex dosimetry, shared uncertainty, multiplicative error

In occupational cohort studies, a panel of experts often creates an exposure matrix or a dosimetry system that estimates dose histories for workers, and then these estimates are used in disease-risk analysis. Errors in the exposure matrix that were shared by time and/or a group of workers were generally ignored. We tested two different methods (Monte-Carlo maximum likelihood and full parametric bootstrap methods) to study the effect of shared uncertainties. The MCML agreed with the uncorrected likelihood ratio test for small additive and small shared multiplicative error distributions. Clear widening of confidence intervals were seen from the MCML and the full parametric bootstrap methods as the shared multiplicative error increased. Although the confidence intervals widened for both methods under the large error model, the range of the confidence intervals disagreed.

The Comparison of Fitted Nonlinear Exposure-Response Relationships in Cox Models Using Smoothing Methods Through Simulations

❖Usha Govindarajulu, Yale University, 2 Livingston St, Apt A32, New Haven, CT 06511, *usha@alum.bu.edu*; Betty Malloy, American University; Bhaswati Ganguli, University of Calcutta; Donna Spiegelman, Harvard School of Public Health; Ellen Eisen, Harvard School of Public Health

Key Words: penalized spline, restricted cubic spline, fractional polynomial, exposure-response, simulations, Cox model

We fit non-linear exposure-response relationships using smoothing techniques in Cox models. Previously, we compared smoothing techniques on actual occupational cohort data, in which time to incidence or mortality is modeled as a function of exposure. We simulated data in order to further compare these smoothing techniques: penalized spline, restricted cubic spline, and fractional polynomial, to the true exposure-response curve, where we know particular characteristics such as the exposure distribution and dose-response curve. We used several measures to compare the fitted curves to the truth: mean-squared error calculations, test for linearity, test for null effect, and % relative bias calculations and concluded that of the various dose-response shapes, the linear dose-response function was the best fitting. Furthermore, the penalized splines were generally closer to the truth.

Recovering the Estimated Slope of an Unobservable Predictor in a Simple Regression

❖Wenyaw Chan, The University of Texas at Houston, 1200 Herman Pressler, Houston, TX 77030, *wenyaw.chan@uth.tmc.edu*; Hung-Wen Yeh, The University of Kansas Medical Center; Elaine Symanski, University of Texas at Houston

Key Words: attenuation, measurement error, regession, surrogate variable

Attenuation of regressor effects in simple regression when a surrogate variable is used as a substitute for a latent predictor has long been a concern in public health. Recognizing that the surrogate variable is an imperfect measure of the true predictor, prior work has focused on estimating attenuation in the slope caused by using the surrogate variable. In addition, there is interest in recovering the estimated slope of the latent predictor. Thus, this study proposes an estimator of the attenuation using conditional expectation of the two slopes estimated from the latent and surrogate predictors given the outcome and surrogate variables. Simulation studies were conducted to examine the performance of this estimator in terms of recovering the estimated slope of the latent predictor. An application of the method is illustrated using data from the occupational health arena.

34 Modeling and Estimation of Disease Prevalence, Incidence, and Spread ●

Section on Statistics in Epidemiology, Section on Health Policy Statistics, Section on Teaching Statistics in the Health Sciences

Sunday, July 29, 2:00 pm–3:50 pm

Informative Allocation To Stabilize MLE in Finite Mixture Models

❖Nong Shang, Centers for Disease Control and Prevention, CDC, Atlanta, GA 30333, *nms6@cdc.gov*

Key Words: Mixture model, EM algorithm, MLE

There is often additional information that may not be easily incorporated into the individual components of a mixture distribution. For instance, extreme values of a test may definitely diagnose or rule out a disease. That is, there may be thresholds of perfect accuracy that may be incorporated into mixture models. Unlike the regular truncated distribution approach, which would involve more complicated computations and may result in less stable MLE, our direct and computationally simple method allocates extreme observations into specific components in the EM algorithm of the MLE. Such method maximizes a slightly different likelihood function. However, it greatly improves the stability of the MLE procedure, resulting in much more accurate estimations. The method has been applied successfully to estimate TB infection prevalence rates. Extensive numerical simulations are also conducted.

An Incidence Model for Capture Recapture

❖Lawrence Lessner, State University of New York at Albany, State University of New York, One University Pl, Rensselaer, NY 12155-3456, *llessner@nycap.rr.com*

Key Words: capture recapture, Poisson point process, decisions

The objective of this work is to present a probability model for the data in a two source capture recapture situation. The model uses nonhomogeneous Poisson point processes and constructs a sequence of point processes that models the construction (incidence of decisions) of the lists. In this model the “capture probabilities” may be time dependent. The resulting data is a 2 by 2 table as usual, whose entries are independent Poisson random variables. The intensities for all point processes are given in closed form.

Spatial Analysis of Breast Cancer Incidence in Rockland County, New York County, and Western Queens County

❖Heather Watson, New York University, 324 E 52nd St Apt 2C, New York, NY 10022, *hnw206@nyu.edu*; Mengling Liu, New York University; Judith D. Goldberg, New York University

Key Words: spatial statistics, small area, aggregate, breast cancer, cluster

At the census tract level, female breast cancer cases are reported from 1990 to 2000 in Rockland County, New York County and Western Queens County in New York State. The United States Census 1990 and 2000 Long Form provided the population at risk and characteristics of interest, such as education and income. Several spatial methods are used to assess breast cancer standardized incidence rates in relation to environmental risk factors, such as distance from local power plants and census tract demographic and community characteristics. The data are analyzed using global and local clustering methods and spatial regression. The results from local clustering analysis show clusters of both increased and decreased incidence in the counties under study.

Hills, Valleys, and Rivers: The Transmission of Raccoon Rabies over an Uneven Landscape

❖David Wheeler, Emory University, 996 Greenwood Ave NE, Apt 6, Atlanta, GA 30306, *dcwheel@sph.emory.edu*; Lance Waller, Emory University

Key Words: GIS, hierarchical models, infectious disease, spatial statistics, landscape epidemiology

Landscape features may serve as either barriers or gateways to the spread of certain infectious diseases, and understanding the way geographic structure impacts disease spread could lead to improved containment strategies. This presentation focuses on modeling the space-time diffusion process of a 1977 raccoon rabies outbreak across several states in the Eastern United States. Specifically, we measure the impact that landscape features, such as mountains and rivers, have on the speed of infectious disease diffusion. This work combines statistical modeling with spatial operations in a geographic information system (GIS) to study disease diffusion. We use a GIS to create landscape feature variables that are input into a hierarchical spatial statistical model to quantify their impacts on transmission time between adjacent counties.

A Simulation Model for Testing the Effect of Interventions During a Pandemic Influenza Outbreak

❖Xiaoping Jin, Centers for Disease Control and Prevention, 1600 Clifton NE, Atlanta, GA 30333, *xjin@cdc.gov*; Hong Zhou, Centers for Disease Control and Prevention; William W. Thompson, Centers for Disease Control and Prevention; Po-Yung Cheng, Centers for Disease Control and Prevention; David K. Shay, Centers for Disease Control and Prevention; Michael Haber, Emory University

Key Words: pandemic, influenza, effectiveness

Pandemic response plans include measures to decrease contact rates between individuals during a pandemic. We used stochastic SIR models to estimate the effectiveness of interventions on pandemic-related illnesses, hospitalizations and deaths. We also compared the effectiveness of allocations of limited vaccine supplies to persons in different age groups. Results suggest that school closings of 1–3 weeks would not substantially affect pandemic-related outcomes. However, if persons who develop symptoms and their household contacts were encouraged to withdraw to their homes, then outcome rates might be reduced by ~50%. Preventing ill long-term care facility residents from making contacts with other residents would reduce morbidity and mortality in this group by ~60%. Finally, vaccination of school children was more effective than vaccination of the elderly in reducing pandemic outcomes.

35 Imputation of Missing Data: Applications and Methods ●

Section on Survey Research Methods, Section on Health Policy Statistics, Biometrics Section

Sunday, July 29, 2:00 pm–3:50 pm

Adjusting the German Labor Force Survey for Under-Reporting by Means of Calibrated Imputation

❖Hans Kiesl, Institute for Employment Research, Regensburger Str 104, Nuremberg, 90478 Germany, *hans.kiesl@iab.de*; Susanne Rässler, Institute for Employment Research

Key Words: calibrated imputation, measurement error, labour force survey, multiple imputation

The German LFS, which is conducted as a CAPI household survey (including proxy interviews), suffers from serious under-reporting of marginal employment. As a result, the number of unemployed is overestimated. To adjust the LFS for this kind of measurement error, we use auxiliary data from administrative files to develop a calibrated imputation algorithm. First, we estimate provisional propensity scores for marginal employment by means of a binary choice model. Then, we change these propensities so that they agree with auxiliary totals from administrative data and at the same time minimize some distance function between the provisional and final propensities. In the last step, we impute a (calibrated) binary variable indicating marginal employment. To get valid variance estimates, multiple imputation is performed. The German NSI intends to use our procedure from May 2007 onward.

Not Just Another Multiple Imputation: Practical Aspects for Highly Correlated Data

❖Lester Curtin, Centers for Disease Control and Prevention, 3311 Toledo Road, Hyattsville, MD 20782, *lrc2@cdc.gov*; Nathaniel Schenker, National Center for Health Statistics; Jeffrey Hughes, Centers for Disease Control and Prevention; Lori Borrund, Centers for Disease Control and Prevention

Key Words: Multiple Imputation, NHANES, Correlated Data

A recent application in the NHANES survey provides interesting insights into practical applications of the multiple imputation technique. The NHANES is a unique survey with self-reported measures, physical measurements, laboratory measurements and an examination known as Dual X-ray absorptbiometry (DXA). DXA measures body fat and bone density for several regions of the body. While some observations seemed to be missing at random, for many observations the missingness was directly related to the outcome measures. This presentation discusses the numerous practical issues that had to be addressed for this unique data set, including the highly correlated data structure, variable transformations, outlier detection, symmetry of imputed measurements, relationships of imputed variables, lower and upper bounds, and the handling of influential observations.

Data Imputation Models for Nontrended Price Data

❖Kennon Copeland, IMS Health, 7216 Ni River Landing, Fredericksburg, VA 22407, *kcopeland@us.imshealth.com*; Boris Brodsky, IMS Health

Key Words: Imputation, Price data, Bias, Variance, Administrative data

Missing data is a common problem for sample surveys that, if ignored, results in increased variance and likely bias for survey estimates. Data values based upon small samples or that have large error may also adversely affect the variance for survey estimates. Imputation, in which values are assigned for missing and "weak" data, is one approach commonly taken to reduce the variance and bias for survey estimates. Imputation for price data is typically carried out through the use of models reflecting change in prices over time. However, if no historical data are available (e.g., initial collection, for rare items), alternative models are required. Approaches for modeling prescription drug prices using related prices, both from sample and from administrative sources, will be presented, performance of the models will be reviewed, and implications for future direction will be discussed.

Weight Adjustments for Fractional Regression Hot Deck Imputation

❖ Min Hui Paik, Iowa State University, 204 Snedecor, Ames, IA 50011, *minhui@iastate.edu*; Michael D. Larsen, Iowa State University

Key Words: missing data, fractional imputation, jackkinfe variance estimation

Fractional regression hot deck imputation (FRHDI), suggested by J. K. Kim, imputes multiple values for each missing dependent variable. The imputed values are equal to the predicted value based on the full observed cases plus multiple random residuals chosen from the set of empirical residuals. Fractional weights are chosen to enable variance estimation and to preserve the correlation among variables. The FRHDI method can be viewed as a special case of fractional hot deck imputation (FHDI). In some circumstances with some starting weight values, existing procedures for computing FHDI weights can produce negative values. We discuss procedures for constructing nonnegative adjusted fractional weights for FRHDI.

Robust Sequential Regression Imputation Method

❖ Irina Bondarenko, University of Michigan, , *ibond@umich.edu*; Trivellore E. Raghunathan, University of Michigan

Key Words: multiple imputation, sequential regression, hierarchical Bayesian, missing data

Sequential Regression Imputation (Raghunathan et al. (2001)) employs a set of regression models to impute missing values in each variable, conditional on all other variables. The imputations, however, are susceptible to model misspecification. We investigate a robust alternative approach. For each variable with missing values, we obtain a predicted value of the variable and a response propensity for missing this variable, conditional on observed and imputed values of other variables. We partition the data set into strata based on the joint distribution of the propensity score and predicted value. A hierarchical Bayesian model across all strata is used to impute the vector of missing values in each stratum. The approach is evaluated using both actual and simulated datasets.

Imputation of Rental Equivalence in the Consumer Expenditure Interview Survey

❖ Barry Steinberg, Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Washington, DC 20212, *steinberg.barry@bls.gov*; Eric Keil, Bureau of Labor Statistics; David Swanson, Bureau of Labor Statistics; Shirley Tsai, Bureau of Labor Statistics; Garner Thesia, Bureau of Labor Statistics

Key Words: Rental Equivalence, Owned Home, Imputation

Data for the rental equivalence of an owned home is collected quarterly by the Consumer Expenditure Interview Survey. The question asked is "If someone were to rent this home today, how much do you think it would rent for monthly, unfurnished and without utilities"? Historically, response rates for this important item have been low. We designed an estimator that will impute rental equivalence values where missings are recorded. After testing several different types of models, we chose a multiple level linear regression model to replace the existing hotdeck method. This paper will focus on a description of the final model that was chosen to be implemented.

36 Sampling Design: Methods, Adaptation, Redesign ●

Section on Survey Research Methods

Sunday, July 29, 2:00 pm–3:50 pm

Multiple Matrix Sampling: A Review

❖ Jeffrey Gonzalez, Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Suite 1950, Washington, DC 20212, *gonzalez.jeffrey@bls.gov*; John L. Eltinge, Bureau of Labor Statistics

Key Words: Split questionnaire, Respondent burden, Nonresponse, Sample survey, Selection probability, Variance estimation

The Consumer Expenditure Quarterly Interview Survey (CEQ) is an ongoing panel survey of US households in which detailed information on an estimated 60 to 70 percent of total expenditures for a consumer unit is collected. The CEQ is generally administered face-to-face and takes about 70 minutes to complete. One proposed method to decrease the length of interview is to use multiple matrix sampling. This involves dividing the questionnaire into sections of questions and then administering these sections to subsamples of the main sample. We provide an overview of the current research on multiple matrix sampling. We review its origins, highlight the fields in which it receives the most application and discuss how it has been applied to problems in surveys. Finally, we discuss the statistical methodology that has been developed for dividing the questionnaire and analyzing the resulting data.

Estimating Population Proportions Using Link-Tracing Sampling Designs with Subsampling and Partial Knowledge on Degree Distribution

❖ Mike Kwanisai, NORC at the University of Chicago, 7146 Talisman Lane, Columbia, 21045, *kwanisai-mike@norc.org*

Key Words: Snowball sampling, Adaptive sampling, Link-tracing, In-degree, Out-degree

Link-tracing sampling designs are used to draw samples from rare, hidden and hard-to-reach populations such as HIV/AIDS patients. In these designs, new subjects are added into the sample by tracing relations from subjects already in the sample. Although convenient, these sampling designs produce unrepresentative samples that make estimation difficult. It is, however, common for investigators to have partial knowledge of some characteristics of the target population, such as the average degree distribution. Estimation procedures that harness and use such partial information produce improved estimates. In this paper, we discuss sampling designs and estimation techniques in which only a fraction of relations are traced to include subjects into the study. Using minimal knowledge about degree distribution, we demonstrate estimation procedures for population proportions.

Some Statistical Issues in Adapting Sample Designs for Surveys Requiring Rapid Response

❖ Allison Deal, The University of North Carolina at Chapel Hill, 730 Martin Luther King, Jr. Blvd., Bolin Creek Center, Suite 104A, Chapel Hill, NC 27599-2400, *allisondeal@unc.edu*; William D. Kalsbeek, The University of North Carolina at Chapel Hill

Key Words: Rapid Response, Overlapping Samples, BRFSS, Sample Size

Special design issues arise when adapting ongoing data gathering systems for quick turn-around assessments, for example during public health emergencies. One of these issues is determining adequate sample size for subgroup estimates. For example, if a comparison between overlapping population subgroups is desired (e.g., when comparing a state coverage rate to one of its region's), special consideration must be given to the overlap when determining sample size requirements. The need to estimate influenza vaccination coverage, in the context of adapting CDC's Behavioral Risk Factor Surveillance System (BRFSS), is used to illustrate these issues and their resolution.

Optimal 2-Stage Sampling Theory for Business Customer Satisfaction Surveys

❖John Rotondo, AT&T Labs - Research, 180 Park Ave, Rm B203, Florham Park, NJ 07932, *rotondo@att.com*

Key Words: 2-stage sampling, optimal sampling, survey sampling, satisfaction surveys

The managers responsible for the selection of methodology used in satisfaction surveys were concerned about potential weaknesses in the sampling method used on larger businesses. The traditional procedure was to survey one knowledgeable high-level manager from each sampled firm. However, on an ad hoc basis, multiple raters were drawn from some of the sampled firms, although no statistical framework was available to guide sampling of multiple raters or analysis of multiple rater data. I present an overview of a 2-stage sampling theory which: 1) accommodates multiple raters from each firm; 2) formally incorporates within-firm variation in satisfaction; 3) enables the designer to solve for the optimal number of raters to sample from each sampled firm in accordance with several alternative statistical optimality criteria; 4) provides a foundation for statistical inference.

An Automated Procedure for Forming Contiguous Sampling Units for Area Probability Samples

❖Bryce Johnson, Westat, 1650 Research Blvd, Rockville, MD 20850, *brycejohnson@westat.com*; Jill M. Montaquila, Westat; Andrew Heller, Westat

Key Words: TIGER, multilevel modeling, segments

In area probability samples, the primary sampling units (PSUs) are often counties or groups of contiguous counties. For household area samples, the secondary sampling units (SSUs, or segments) are often constructed using Census blocks, block groups, or tracts. Algorithms to form segments by grouping areas that are adjacent in numbering schemes (e.g., adjacent block numbers within tract, in a sorted list) sometimes result in discontiguous segments. This may be a problem if area measures (e.g., environmental data) are to be obtained or if multilevel modeling is to be used for analysis. In this paper we describe a study that motivated the development of an algorithm to optimize the configuration of segments based on Census blocks. The algorithm is described and evaluated, and ideas for enhancements are discussed.

Selection of Small and Large Schools in State and County School Surveys

❖Tonja M. Kyle, Macro International, 11785 Beltsville Drive, Suite 300, Calverton, MD 20705, *tonja.m.kyle@orcmacro.com*; Pedro J. Saavedra, Macro International; James G. Ross, Macro International

Key Words: PPS sampling, minimum replacement, stratification, design effect, cluster sample, intra-class correlation

The most common method when conducting state or county school surveys is to first select schools with probabilities proportional to size using enrollment as the measure of size, and then sample the same number of students from every school. Ideally, this would result in the same probability of selection for every student in the frame. One difficulty is that in both state and county surveys there are usually a number of schools with enrollments below the number of students targeted for each participating school. A second difficulty, primarily present in county surveys, is that there are often schools where the calculated probability of selection exceeds 1.0. This paper explores different ways of handling the above problems and their impact on design effect using simulations from publicly available school data. Variables with low and high intra-class correlations are used.

Reduction of Sample Areas in the Consumer Price Index and Consumer Expenditure Survey Designs

❖Lawrence Ernst, Bureau of Labor Statistics, 2 Massachusetts Ave NE Rm 1950, Washington, DC 20212, *ernst.lawrence@bls.gov*; William H. Johnson, Bureau of Labor Statistics; William E. Larson, Bureau of Labor Statistics

Key Words: PSUs, maximization of overlap, PSU weights

The Consumer Price Index (CPI) and the Consumer Expenditure Survey (CE) are surveys with multistage designs, revised every 10 years. The first stage CPI and CE samples include a set of areas (PSUs) selected from the set of U.S Core Based Statistical Areas. CE additionally selects a set of PSUs to represent the rest of the nation. After selecting the original sample of PSUs, a reduction was considered for budgetary reasons, a reduction implemented in CE only. In this paper we describe: the details of the reduction process used and alternative approaches; the adjustment of the PSU weights resulting from the reduction, which was complicated by the use of a maximization of overlap procedure in the original selection of the new sample PSUs; possible improvements to the overlap procedure in the next redesign; and the impact of having different sets of CPI and CE sample PSUs.

37 Introductory Overview Lecture 1: Adaptive Designs and Other Emerging Issues in Clinical Trials

The ASA, ENAR, WNAR, IMS, SSC, Committee on Applied Statisticians

Sunday, July 29, 4:00 pm–5:50 pm

Designing Clinical Trials To Evaluate Dynamic Treatment Regimes

❖Peter F. Thall, The University of Texas M.D. Anderson Cancer Center, Dept. of Biostatistics, Box 447, 1515 Holcombe Blvd., Houston, TX 77030, *rex@mdanderson.org*

Key Words: Adaptive design, Bayesian statistics, Cancer therapy, Clinical trial, Treatment policy

A dynamic treatment regime, also known as a treatment policy or multistage treatment strategy, is a set of formal rules for what physicians do routinely when treating a patient in multiple stages. The physician makes a baseline diagnosis by evaluating the patient's signs and symptoms, chooses and administers an initial treatment, observes the patient's outcomes, and uses all of this information to choose the next treatment. This process is repeated until criteria for stopping the patient's therapy are met. I will describe some cancer clinical trials involving multi-stage dynamic treatment regimes that I have designed using both Bayesian and frequentist methods. I also will discuss some of the practicalities, computational requirements, and statistical issues arising in the design and conduct of clinical trials involving dynamic treatment regimes.

When Would I Consider Adaptive Design for Late-Stage Development?

❖ Keaven M. Anderson, Merck & Co., Inc., Clinical Biostatistics and Research Decision Sciences, PO Box 4, West Point, PA 19486-0004, *keaven_anderson@merck.com*

Key Words: adaptive design, clinical trials, Phase II/III

This talk is based on primarily on past experiences for specific trials where adaptation was found to be useful. A primary consideration is whether or not there is time to do adaptation given the speed of enrollment and delays due to patient follow-up, data entry, data cleaning, analysis and decision making. How much information adaptation reveals and its affect on regulatory acceptance must be considered. Specific adaptations will include fully sequential analysis for safety issues, group sequential design and use of information-based timing of analyses, adaptive sample size re-estimation, dropping of arms and efficiency issues. Another area that will be discussed is examples from Phase II/III combination studies in oncology. In addition to adapting within a trial, using limited interim information to make other decisions related to product development will be discussed.

38 Curve Forecasting in Modern Business ● ✪

Business and Economics Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data

❖ Jonathan Stroud, University of Pennsylvania, Wharton School, Philadelphia, PA 19104, *stroud@wharton.upenn.edu*; Jonathan Weinberg, University of Pennsylvania; Lawrence D. Brown, University of Pennsylvania

Key Words: Autoregressive models, Bayesian forecasting, Call centers, Cubic smoothing spline, Poisson process, Sequential Monte Carlo

A call center is a centralized hub where customer and other telephone calls are dealt with by an organization. In today's economy, they have become the primary point of contact between customers and businesses. Accurate prediction of the call arrival rate is therefore indispensable for call center practitioners to staff their call center efficiently and cost effectively. This article proposes a multiplicative model for modeling and forecasting within-day arrival rates to a US commercial bank's call center. Markov chain and sequential Monte Carlo methods are used to estimate both latent states and model parameters. The calibration of these predictive densities is evaluated through probability integral transforms. Furthermore, we provide one-day-ahead forecast comparisons with classical statistical models. Our predictions show improvements of up to 25% over these standards.

Interday Forecasting and Intraday Updating of Call Center Arrivals

❖ Haipeng Shen, The University of North Carolina at Chapel Hill, 304 Smith Building, Chapel Hill, NC 27599, *haipeng@email.unc.edu*; Jianhua Z. Huang, Texas A&M University

Key Words: dimension reduction, dynamic forecast updating, principal component analysis, penalized least squares, singular value decomposition, vector time series

Accurate forecasting of call arrivals is critical for staffing and scheduling of a call center. We develop methods for interday and intraday forecasting of incoming call volumes. Our approach is to treat the intraday call volume profiles as a high-dimensional vector time series. We propose to first reduce the dimensionality by singular value decomposition of the matrix of historical intraday profiles and then apply time series and regression techniques. Both interday dynamics and intraday patterns of call arrivals are taken into account by our approach. Distributional forecasts are also developed. Our methods are data-driven, and appear to be robust against model assumptions in simulation studies. They are shown to be very competitive against existing approaches in out-of-sample forecast exercises using real data. Our methods are computationally fast for real time dynamic forecasting.

Bayesian Modeling and Forecasting of Intraday Electricity Load and Spot Prices

❖ Michael S. Smith, University of Melbourne, Melbourne Business School, 200 Leicester Street, Carlton, International VIC 3053 Australia, *mike.smith@mbs.edu*

Key Words: MCMC, Multi-equation models, Multivariate SV, Covariance Selection, Applied Bayesian smoothing, Forecasting

With the advent of wholesale electricity markets there has been renewed focus on intraday forecasting of both electricity demand (or load) and spot prices. This presentation discusses a variety of approaches for modeling both electricity load and spot prices. In particular, the applicability of contemporary Bayesian smoothing, multiequation modeling and multivariate time series for these problems is examined. The models are large and highly parameterized, but can still be estimated effectively using Markov chain Monte Carlo. The role of Bayesian model and variable selection methodology in such models is examined, as well as the ability to obtain the full predictive distribution of key quantities of interest. While the focus is primarily on forecasting intra-day load, the case of spot prices is also examined. The models and methods are illustrated using Australian intraday data.

39 Design and Analysis for Genome-Wide Association Studies ● ✪

ENAR, Section on Statistics in Epidemiology, Biometrics Section, WNAR

Sunday, July 29, 4:00 pm–5:50 pm

Statistical Issues Arising from Analysis of Genome-Wide Case-Control Association Studies

❖ Gang Zheng, National Heart, Lung, and Blood Institute, 6701 Rockledge Drive, MSC 7938, Bethesda, MD 20892, *zhengg@nhlbi.nih.gov*; Jungnam Joo, National Heart, Lung, and Blood Institute; Colin O. Wu, National Heart, Lung, and Blood Institute; Nancy L. Geller, National Heart, Lung, and Blood Institute

Key Words: Quality control, Multiple testing, Population structure, Replication, Robust test

We discuss several statistical issues arising from our analysis of a genome-wide association study (GWAS) using over 100K SNPs. In quality control, we discuss the rational of testing HWE and a possible trade-off between high per SNP call rate and high per SNP error rate. In genome-wide scans (GWS), at least three common, simple test statistics can be used. We propose a robust tag statistic. Multiple testing is a serious issue in analysis of GWAS, starting even with quality control. We discuss when one should correct the multiple testing. In using genomic control or delta centralization to correct for cryptic relatedness and population stratification, null loci are used. Issues of how to choose null loci will be discussed. Finally, we discuss issues in replicating results from a GWS, including sample size calculations and how to set the significance level for replications.

Multistage Sampling for Genetic Studies

❖Robert C. Elston, Case Western Reserve University, Department of Epidemiology and Biostatistics, University School of Medicine, Wolstein Research Building, 2103 Cornell Road, Cleveland, OH 44106-7281, *rce@darwin.case.edu*

Key Words: Association, DNA pooling, Genome-wide, Linkage analysis, Replication, Two-stage

In the past, to study Mendelian diseases, segregating families have been ascertained for segregation analysis, followed by linkage analysis. This would then be followed by association studies. Recently, for more complex diseases, the initial sampling has been for a genome-wide linkage analysis to identify candidate regions for follow-up with association studies. We now have the ability to conduct genome-wide association studies using 100K to 500K diallelic genetic markers. For such studies we focus on efficient two-stage association sampling designs, which can retain nearly optimal statistical power at about half the genotyping cost. Similarly, beginning an association study by genotyping pooled samples may also be a viable option if the cost of accurately pooling DNA samples outweighs genotyping costs.

Testing Strategies for Genome-Wide Association Studies: Using the Same Dataset for Screening and Replication

❖Christoph Lange, Harvard School of Public Health, Department of Biostatistics, 665 Huntington Ave, Boston, MA 02115, *clange@hsph.harvard.edu*

Key Words: Family-based studies, Genome-wide association, Multiple testing problem, Replication, Screening, Two-stage

Genome-wide association study (GWAS) is now a reality. Their success will depend on the efficient handling of the statistical challenges posed by GWAS. GWAS must translate the markedly increased amount of SNP-information into increased statistical power. For the number of statistical tests computed in a GWAS, standard statistical methods for handling the multiple testing problem (MTP) are too conservative. For QTL, VanSteen et al (2005) proposed a new approach that bypasses the MTP in family-based studies. The approach consists of a screening step and a testing step, which handle the MTP all in the same data set. It outperforms standard methodology in terms of statistical power (e.g., it addresses the MTP and the replication problem both at the same time). We will review the VanSteen approach, and discuss extensions to dichotomous traits in trio designs and to case/control design.

Interactions in Whole-Genome Association Studies

❖Josephine Hoh, Yale University, 60 College Street, New Haven, CT 06510-3210, *josephine.hoh@yale.edu*

Key Words: genome-wide, epistasis, association

Whole genome association mapping has been proven to be efficient in searching susceptibility genes for human illnesses. Data in this type of studies are enormous. While association tests generally look for a main effect, interactions between loci (epistasis) also contribute to phenotypic variation. This presentation will focus on SNP interactions in disease association. Two-stage approach has been developed. We applied our methods to the real data containing 100K and 300K SNPs. At the first stage we directly screen interactions without main effects being specified. Based on the scales of relative risks, we identify those pairs of SNPs with crossover responsive curves under all possible disease inheritance modes. At the second stage, we evaluate the significance of the interactions in disease association and also try to construct the hypothetical genetic networks.

40 First Hitting Time Models and Threshold Regression ● ✪

SSC, Biometrics Section

Sunday, July 29, 4:00 pm–5:50 pm

What Hides Behind the Data?

❖Odd O. Aalen, University of Oslo, Dept. of Biostatistics, PO Box 1122 Blindern, Oslo, N-0317 Norway, *o.o.aalen@medisin.uio.no*

Key Words: survival analysis, shape of hazard rate, quasi-stationarity, underlying process, hitting-time models, non-proportional hazards

In survival analysis the focus is usually on the mere occurrence of events. Not much emphasis is placed on understanding the processes leading up to these events. The simple reason for this is that these processes are usually unobserved. However, one may consider the structure of possible underlying processes and draw some general conclusions from this. For instance, a practically important issue is why hazard rates assume various specific shapes, e.g. why do we often see hazard rates that first increase and then decrease? We shall show the usefulness of the concept of quasi-stationarity. These are stationary distributions that arise in stochastic processes where probability mass is continuously being lost to some set of absorbing states. Due to this leaking of probability mass, the limiting distribution is just stationary in a conditional sense that is, conditioned on nonabsorption.

A Class of Simple Lifetime Regression Models Resulting from First-Hitting Times

❖Thierry Duchesne, Laval University, Departement de mathematiques & statistique, Pavillon Alexandre-Vachon, Quebec, QC G1K 7P4 Canada, *duchesne@mat.ulaval.ca*

Key Words: lifetime regression, threshold regression, collapsible model, time-varying covariate

Oakes (1995), Kordonsky and Gertsbakh (1997) and Duchesne and Lawless (2000) consider a class of simple regression models for lifetime given the history of an increasing covariate that can be viewed as an alternate measure of age. These models are appealing due to their interpretability and potential for non-parametric inference, and they fit several types of datasets quite well; Oakes (1995) uses them to model the lifetime of miners exposed to asbestos dust and in an analysis of the Channing House retirement center data, whereas Kordonsky and Gertsbakh (1997) and Duchesne and Lawless (2000) use them in reliability data analyses. In this talk, we will see how these models can be obtained as the distribution of the first time that a wear (or health) process that is driven by usage (or exposure) hits a threshold. Potential for specific applications will also be discussed.

Longitudinal Data in Threshold Regression: Implementation and Relation to Cox Regression with Time-Varying Covariates

❖Mei-Ling Ting Lee, The Ohio State University, B-122 Starling-Loving Hall, Biostatistics Division, 320 West 10th Avenue, Columbus, OH 43210, *meiling@sph.osu.edu*; George A. Whitmore, McGill University

Key Words: survival analysis, Cox regression, longitudinal data, wiener process, boundary crossing, time-varying covariates

Longitudinal survival data pose an interesting challenge. Lee and Whitmore (2007) review a new regression methodology for survival data referred to as threshold regression. The methodology is based on the concept that degradation of an item follows a stochastic process and failure occurs when the process first reaches a failure state or threshold. Breaking longitudinal records into series of single records is one strategy that has been proposed. This study looks at the formal conditions that must hold for this uncoupling procedure to be valid. The conditions are examined in terms of both theory and practical application. The uncoupling procedure modifies the time scale for the analysis and can be used in conjunction with an operational time scale. We show that the Cox proportional hazards regression model with time-varying covariates is a special semi-parametric version of the model.

41 The Role of Statistical Standards in Federal Surveys ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Evaluation of Prospective Costs and Benefits of Standards for Statistical Programs

❖John L. Eltinge, Bureau of Labor Statistics, OSMR, PSB 1950, 2 Massachusetts Avenue NE, Washington, DC 20212, *Eltinge.John@bls.gov*

Key Words: Adoption and diffusion of technology, Constraints, Local and global optimization, Risk management, Satisficing, Transparency

Standards often have an important role in the development, adoption and diffusion of many types of technology. This paper considers statistical survey methodology as a form of technology, and uses the resulting conceptual framework to explore several ways in which to evaluate the prospective costs and benefits of standards for statistical programs. The conceptual framework places primary emphasis on types of standards; methods for calibration; methods for application and enforcement; special issues in statistical standards; and special issues for government-sponsored surveys. This framework leads to discussion of prospective benefits, costs and risks of standards. The paper closes with comments on practical implications of these general ideas for development, implementation and enforcement of standards; training and management of statistical program personnel; and communication.

Developing and Implementing Data Quality Standards at the U.S. Census Bureau

Pamela D. McGovern, U.S. Census Bureau; ❖David C. Whitford, U.S. Census Bureau, 4700 Silver Hill Rd., Washington, DC 20233, *david.c.whitford@census.gov*; John M. Bushery, U.S. Census Bureau

Key Words: quality, standards, framework

To ensure consistent quality across all its programs and products, the Census Bureau began developing formal data quality standards in 2001. As of April 2007, eleven quality standards have been published. In 2006, the Census Bureau accelerated its efforts to develop a comprehensive and cohesive set of data quality standards. These standards will customize the Office of Management and Budget's (OMB) Standards and Guidelines for Statistical Surveys to reflect the issues affecting the Census Bureau's processes and products. This paper discusses the Census Bureau's efforts to develop and implement data quality standards. It describes the development approach, which includes defining data quality, constructing a quality framework, and conducting focus groups to gather program area input. This paper also discusses the challenges encountered in developing and implementing these standards.

Lessons Learned Using Statistical Standards To Improve/Ensure Data Quality

❖Marilyn Seastrom, National Center for Education Statistics, 1990 K Street NW, Room 9047, Washington, DC 20006, *Marilyn.Seastrom@ed.gov*

Key Words: Standards, data quality, nonresponse bias, response rates, timeliness, imputations

NCES issued revised Statistical Standards in 2002 that touch on many of the topics covered in the 2006 OMB Statistical Standards and Guidelines for Statistical Surveys. One component of the Statistical Standards Program at NCES is the monitoring of the implementation of the NCES standards. These activities provide the basis for monitoring the timeliness of NCES data collections, tracking trends in unit, overall, and item response rates over time; monitoring whether the use of incentives impacts trends in response rates; tracking the number of NCES data collections requiring unit level nonresponse bias analyses, whether the results of those analyses are used in nonresponse weight adjustments, and the impact of those adjustments on the amount of measurable bias; and evaluating when imputations are used to account for item nonresponse.

42 Collaboration Between Academia, Industry, and Government in Statistical Sciences ●

SPAIG Committee, Section on Statisticians in Defense and National Security, ENAR

Sunday, July 29, 4:00 pm–5:50 pm

The 2007 SPAIG Salary Survey

❖Michael D. Larsen, Iowa State University, 220 Snedecor Hall, Department of Statistics, Ames, IA 50011, *larsen@iastate.edu*; Dianne Anderson, Iowa State University; Janice Larson, Iowa State University

Key Words: Response mode, Stratified sample, Survey experiment, Telephone survey, Internet survey, Salary survey

The Center for Survey Statistics and Methodology (CSSM) at Iowa State University will conduct the 2007 SPAIG salary survey of statisticians in business, industry, and government in the spring of 2007. Building on previous improvements in the sampling design and questionnaire, CSSM anticipates continued high response rates and high quality data. As part of the 2007 survey, a response-mode experiment will be conducted to compare the impact of mail versus email contact. This talk will summarize key results of both the salary survey and the planned statistical investigations.

Student Recruitment Database

❖Robert G. Wilkinson, The Lubrizol Corporation, 29400 Lakeland Boulevard, Mail Drop 152B, Wickliffe, OH 44092-2298, *rgw@lubrizol.com*

Key Words: Recruitment, Database, SPAIG

As part of a SPAIG objective to facilitate the use of statisticians in industry, an initiative was kicked off to improve the process that Business/Industry/Government (B/I/G) organizations use to recruit graduate statistics students. This talk will discuss a prototype that resulted from this initiative (i.e., a searchable data-driven system that allows B/I/G organizations to locate new graduates whose qualifications and professional interests best match the intern and permanent career opportunities within their organizations). The future of the prototype will also be discussed.

The WPI-NCHS Partnership

❖ Balgobin Nandram, Worcester Polytechnic Institute, Mathematical Sciences, 100 Institute Road, Worcester, MA 01609, *balnan@wpi.edu*

Key Words: Collaboration, Health data, Internship, Journal articles, MS theses, Research

Over the past 8–10 years, there has been an excellent working relationship between Worcester Polytechnic Institute (WPI) and the National Center for Health Statistics (NCHS). This relationship has been fueled by collaboration between WPI faculty and students and scientists at the NCHS. Most of these students were Master's students, and they worked for a major part of their theses at the NCHS. As a result, many articles have been published in peer reviewed journals, with collaboration from scientists at the NCHS, WPI faculty and students. In this talk, many different features of the collaboration in the WPI-NCHS partnership are discussed.

How to Compete for the Academe, Industry, and Government (SPAIG) Award?

❖ Jai W. Choi, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *jwc7@cdc.gov*; Robert Starbuck, Wyeth

Key Words: SPAIG, Institution, Partnership, Award, Application

Every year, Statistics Partnerships among Academe, Industry and Government (SPAIG) committee select a winning partnership for SPAIG Award. This award is a major ASA award presented at the annual JSM, and it is not given to individuals but to institutions whose partnership produced successful outcomes beneficial to both parties. However, there are a limited number of applicants because they do not know how easy it is to apply or are not aware of the award. The purpose of this talk is to study the history and description of the award and to encourage possible candidates to apply for the SPAIG Award. I will talk on how to print out the application form and how to write supporting letters to win the award in 2007, showing some examples of the past winners.

43

Rietz Lecture

IMS

Sunday, July 29, 4:00 pm–5:50 pm

Statistical Problems of Gene Mapping

❖ David Siegmund, Stanford University, Department of Statistics, Stanford, CA 94305-4020, *dos@stat.stanford.edu*

Gene mapping attempts to discover the genomic regions that contain genes contributing to specific phenotypes. It is used in experimental genetics to improve the hardiness or productivity of domestic animals or plants of agricultural value, explore basic mechanisms of inheritance, and/or study animal models of human inheritance. In human populations, it is used as a first step to identify genes associated with human health and disease. In this talk, I will discuss statistical issues of gene mapping, from the fundamental contribution of Fisher (1918) to the problems of today. Central issues involve models connecting genes and environment to phenotypes, statistical irregularities in the analysis of these models, multiple comparisons involved in testing many genetic markers for correlation with a phenotype or phenotypes, computational problems, and the role of population history.

44

Challenge Issues on Longitudinal Biomedical Data Analysis ● ✪

Biometrics Section, ENAR, Section on Statistics in Epidemiology, Section on Health Policy Statistics, WNAR

Sunday, July 29, 4:00 pm–5:50 pm

Does Mis-specification of the Random Effects Distribution Affect Predictions of Random Effects?

❖ Charles E. McCulloch, University of California, San Francisco, 185 Berry Street, Suite 5700, San Francisco, CA 94107, *chuck@biostat.ucsf.edu*; John Neuhaus, University of California, San Francisco

Key Words: mixed models, misspecification, generalized linear models

Statistical models that include random effects are commonly used to analyze longitudinal and clustered data. Predicted values of the random effects are often used, e.g., in profiling of physician or hospital performance. In typical applications, the data analyst specifies a parametric distribution for the random effects (often Gaussian) although there is little information available to guide this choice. Whether inferences about regression parameters are sensitive to this specification is of considerable debate in the literature. However, there has been little work on whether the misspecification affects prediction of random effects. Through theory, simulation and an example, we show that misspecification can have a moderate impact on predictions of random effects and develop simple ways to diagnose such sensitivity.

Modeling of Mean-Covariance Structures in Generalized Estimating Equations for Longitudinal Data

❖ Jianxin Pan, University of Manchester, School of Mathematics, University of Manchester, PO Box 88, Sackville Street, Manchester, M60 1QD United Kingdom, *jianxin.pan@manchester.ac.uk*

Key Words: Cholesky decomposition, Efficiency, Generalized estimating equation, Longitudinal data, Misspecification of covariance structure, Modelling of mean-covariance structures

When used for modeling longitudinal data generalized estimating equations specify a working structure for the within-subject covariance matrices, aiming to produce efficient parameter estimators. However, misspecification of the working covariance structure may lead to a large loss of efficiency of the estimators of the mean parameters. In this talk I will introduce an approach for joint modeling of the mean and covariance structures for longitudinal data within the framework of generalized estimating equations. The resulting estimators for the mean and covariance parameters are shown to be consistent and asymptotically Normally distributed. Real data analysis and simulation studies show that the proposed approach produces efficient estimators for both the mean and covariance parameters.

Finite Sample Bias Corrections to Sandwich Covariance Estimators for Longitudinal and Clustered Data

❖ John Preisser, The University of North Carolina at Chapel Hill, Department of Biostatistics, Chapel Hill, NC 27599-7420, *john_preisser@unc.edu*; Bing Lu, Brown University; Bahjat Qaqish, The University of North Carolina at Chapel Hill

Key Words: Cluster trials, Correlated binary data, Generalized estimating equations, Intraclass correlation, Sandwich estimator

Empirical sandwich covariance estimators are widely used in the regression analysis of biomedical longitudinal and clustered data owing to the desire to avoid modeling complex correlation structures often considered secondary to the parameters of interest. Unfortunately, when the number of clusters, or subjects in a longitudinal study, is small, sandwich estimators may underestimate the true variances of estimated regression coefficients and lead to undercoverage of confidence intervals. We review some bias corrections proposed for the first-order GEE sandwich estimator, and show that similar adjustments for modeling intracluster association lead to improved finite sample properties. The proposed methods are illustrated using biobehavioral data from a nested pretest-posttest cross-sectional cluster intervention trial on reducing underage drinking.

Analysis of Clustered Longitudinal Data with Applications to Clinical Dental Research

❖Brian G. Leroux, University of Washington, Departments of Biostatistics, Dental Public Health, Box 357475, Seattle, WA 98195-7475, *leroux@u.washington.edu*

Key Words: clustered, longitudinal, multilevel, regression, estimating equation

Clinical dental data typically has a multilevel structure with clustering of teeth within patients and multiple observations over time for each tooth. The multilevel structure and the large number of observations per patient presents challenges for formulating valid and efficient methods for fitting regression models. Generalized Estimating Equations provides a useful framework for estimation but presents challenges in specifying a working correlation structure that will yield efficient estimates and accommodate missing or censored outcomes. We consider a new approach for modeling correlation structure of clustered longitudinal data and apply it to estimating equations for fitting regression models. Using computer simulation and application to real clinical dental data, the new approach is compared to standard GEE methods in terms of precision and susceptibility to missing data bias.

45 Probing the Universe with Nonparametric Methods

IMS, Biometrics Section, Section on Nonparametric Statistics

Sunday, July 29, 4:00 pm–5:50 pm

Statistical Issues in Detecting Gamma-Ray Pulsars

❖John Rice, University of California, Berkeley, Department of Statistics, Berkeley, CA 94720-3860, *rice@stat.berkeley.edu*

Key Words: TBA

I will discuss our on-going work on methodology for detecting gamma- ray pulsars from data consisting of the arrival times and energies of individual photons. I will discuss a family of detection procedures—weighted quadratic test statistics arising from a score test. The likelihood function models the photon arrival process as a mixture from background and source and takes into account the point spread function and the energies of the photons with the result that a weight is associated with each photon. The power of the test can be analyzed rather explicitly. On the computational side, searching over a range of frequency and frequency drift can be a daunting task, even for a record consisting of only a thousand events or so, and in some situations a tradeoff between power and computation must be made.

Semiparametric Bivariate Density Estimation with Irregularly Truncated Data

❖Chad Schafer, Carnegie Mellon University, Department of Statistics, 5000 Forbes Avenue, 132 Baker Hall, Pittsburgh, PA 15213, *cschafer@stat.cmu.edu*

Key Words: semiparametric method, truncated data, astronomy

This work develops an estimator for the bivariate density given a sample of data truncated to a nonrectangular region. Such inference problems occur in various fields; the motivating application here was a problem in astronomy. The approach is semiparametric, combining a nonparametric local likelihood density estimator with a simple parametric form to account for the dependence of the two random variables. Large sample theory for M-estimators is utilized to approximate the distribution for the estimator. Results are described from the analysis of data from the measurements of quasars.

Clustering of Absorption Systems

❖Ji Meng Loh, Columbia University, Department of Statistics, New York, NY , *meng@stat.columbia.edu*

Key Words: TBA

Absorption systems or absorbers are observed on lines of sight from the earth to distant quasars and are believed to be gas clouds near galaxies that are too far away to be easily detected in large surveys. The absorbers trace the distribution of nonluminous matter around these galaxies. Of scientific interest is the clustering of absorbers and how this compares with that found in galaxy surveys. We discuss how second and third order moment properties of absorbers can be estimated and how this can be used to investigate whether absorber data show evidence of being distributed along filaments and sheets like galaxies are.

46 Gerontologic Biostatistics ●✪

Section on Statistics in Epidemiology, Biometrics Section, WNAR

Sunday, July 29, 4:00 pm–5:50 pm

Integrating Qualitative and Quantitative Information in Clinical Epidemiologic Research with Older Populations

❖Peter H. Van Ness, Yale University, Program on Aging, School of Medicine, 1 Church Street / 7th Fllor, New Haven, CT 06510, *peter.vanness@yale.edu*

Key Words: gerontologic biostatistics, qualitative information, scale construction, calibration, elicited priors, Bayesian methods

Qualitative studies figure prominently in clinical aging research because the experience of aging has ineluctably subjective aspects and because illness in the context of imminent death raises important legal, ethical, and spiritual issues. Yet the predominant inferential methodology of biomedical research among aging and older populations is quantitative. Thus, integrating qualitative and quantitative information is an important challenge for gerontologic biostatisticians. Three integration strategies are reviewed: 1) Qualitative information from open-ended questioning is used to identify pertinent categories for scale construction; 2) Qualitative questions are included in a quantitative study as a means for qualitative calibration of regression models; and 3) Quantitative information reflecting “expert opinion” is used as elicited priors in Bayesian analyses.

Meta-Analysis at the Individual Level: Synthesis of Inferences Across Multiple Gerontological Databases

❖Karen Bandeen-Roche, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, 615 N Wolfe Street, Baltimore, MD 21205, *kbandeen@jhsph.edu*

Key Words: aging, selection, differential measurement, bias

Multiple cohort studies have studied health declines in older adults. Health aspects assessed overlap, hence combining individual-level information across the studies for addressing research questions on elders' health has promise. We propose methodology to heighten accuracy and precision of inferences in analyses that combine so. We address proper accounting for stochastic heterogeneity across studies, as well as two challenges that may be primarily biasing rather than variability heightening: the least or most healthy are often screened out, thus target populations may differ considerably across studies; and key outcomes may be subject to appreciably differential measurement. The project exemplifies multiplicity of challenges in aging statistics and aims to improve accuracy and precision of findings obtainable through the combined use of multiple data sets on aging.

Multiple Trajectories in Aging Research

❖Carl F. Pieper, Duke University, Box 3003 Center on Aging, Medical Center, Durham, NC 27710, *carl.pieper@duke.edu*

Key Words: gerontology, mixed model, latent class, multivariate

Under the usual mixed model design, outcomes are often analyzed 'univariately.' In aging research, multiple systems are often studied simultaneously (e.g., ADLs, IADLs, cognitive impairment, physical performance, quality of life). Change over time or trajectories across these multiple systems are then correlated. Snijders (1999) and Thiebaut (2002) demonstrate a bivariate hierarchical model to simultaneously model 2 outcomes. The two level mixed model will be extended to see if, in longitudinal studies, (1) the 'trajectories' from the multiple outcomes are correlated, and (2) if classes of trajectories can be defined. Using data from the Duke EPESE study, change in measures of ADLs, IADLs, Cognitive Impairment, and Quality of Life will be estimated. Models correlating the empirical Bayes parameters of the trajectories estimated univariately will be contrasted to multivariate models.

47 Land Grant University Ethics (LANGURE): The Ethical Use of Statistics in Research ● ✪

Section on Statistical Consulting, Section on Statistical Education

Sunday, July 29, 4:00 pm–5:50 pm

The Ethical Use of Statistics in Research

❖George P. McCabe, Purdue University, 150 N University Street, West Lafayette, IN 47907-2067, *mccabe@stat.purdue.edu*

Key Words: ethics, consulting

As a consultant, the statistician is sometimes viewed as a gatekeeper with the primary responsibility of preventing researchers from violating statistical principles. We do not want to do anything wrong. With the trend toward collaborative research that is happening in many settings, however, the role of statistician has changed. We have greater responsibilities and we need to deal with situations where there is no clear right or wrong decision. The ethical implications of this expanded role will be explored and illustrated with examples. Transparency in decision making is offered as a guiding principle for ethical behavior in ambiguous situations.

Training University Students in Ethical Use of Statistics as Part of the LANGURE Project

❖Larry A. Nelson, North Carolina State University, Campus Box 7608, Raleigh, NC 27695/7608, *larry_nelson@ncsu.edu*

Key Words: Ethics in Research, Ethical Use of Statistics, Misuse, LANGURE

The need for an Ethics in Research course gave rise to a project called LANGURE involving a consortium of eight U. S. universities. The resulting one-hour course will be taught at several Land Grant and related institutions. In some institutions, it will be required of all graduate students. One important course module will be Ethical Use of Statistics. Statisticians and users of Statistics from the eight universities are jointly involved in developing the content of the course. This content will focus on types of ethical problems in statistical usage such as misuse, sloppiness and outright fraud. Case study examples of each of these problem types will be presented in the course. In addition to planning and conducting the experiments or surveys, emphasis will also be given to the analysis and reporting phases. Module access will be available on the NCSU Graduate School web site.

Land Grant University Ethics (LANGURE): The Ethical Use of Statistics in Research

❖Michael Crotty, SAS Institute Inc., SAS Campus Drive, S4 4089, Cary, NC 27513, *michael.crotty@sas.com*

Graduate students conducting research at land grant universities benefit from training in research ethics, particularly the ethical use of statistics in their research. The ethical use of statistics includes involving a statistician from the design phase of a study through to the analysis and presentation of results. A research ethics course has been designed for graduate research students to learn about conducting ethical research. We focus on the statistics module developed for this course.

48 Design- and Model-Based Sampling and Inference ●

Section on Statistics and the Environment, Section on Statistics in Epidemiology, Biometrics Section, General Methodology, Section on Survey Research Methods

Sunday, July 29, 4:00 pm–5:50 pm

Design-Based, Model-Based, and Model-Assisted Sampling, and Inference in Natural Resources

❖Mary Christman, University of Florida, Department of Statistics, PO Box 110339, Gainesville, FL 32611-0339, *mcxman@ufl.edu*

Key Words: environmental sampling, optimal sampling strategies

The differences in design-based versus model-based versus model-assisted sampling strategies and inferential procedures are often not well-understood when applied to environmental and ecological research. In this presentation we describe, compare, and contrast the three approaches in several ways, including identifying strategies for sampling design, choice of estimation or inference procedures, and examining the conditions under which the approaches are optimal or at least appropriate. In addition, there will be discussion of the potential for misuse of the approaches. Several examples are given to demonstrate similarities and differences among the three approaches.

Optimal Spatial Sampling: Probability-Based or Model-Based?

❖Don L. Stevens, Jr., Oregon State University, Statistics Department, 44 Kidder Hall, Corvallis, OR 97333, *stevens@stat.oregonstate.edu*

Key Words: environmental sampling, optimal design, spatial simulated annealing

Assessing a natural resource can use either model-based or design-based methodology. A spatial model may have to estimate the semi-variogram in order to predict the mean value of realized population. Good semi-variogram estimates require points with small separations. In contrast, design-based estimates are generally more precise with maximal separation. Design optimality can be judged by different criteria, even with the same objective, so that the "optimal" designs may be very different. Model-based assessment can be applied regardless of how samples were picked, but design-based methods can be used only with a probability sample. In some situations, e.g., litigation, model-free estimates can be advantageous. I contrast model-optimal and probability-optimal sampling. I also explore the notion that probability designs can be near-optimal judged by model-based and design-based criteria.

Design-Based and Model-Based Issues in Natural Resource Inventories

❖Ronald E. McRoberts, USDA Forest Service, Northern Research Station, 1992 Folwell Avenue, Saint Paul, MN 55108, *rmcroberts@fs.fed.us*

Key Words: Variance estimators, Forest inventory, Small area estimation

Traditionally, large area natural resource inventories have relied on probability-based sampling and design-based inferences. However, the advent of moderate resolution satellite imagery and ground-based spatial data as sources of ancillary data, geographic information systems for combining data from these multiple sources, and fast computer processing systems have made model-based approaches to inference more attractive. Unfortunately, differences in the conceptual underpinnings of design-based and model-based approaches are often not well-understood within these natural resource inventory communities. The emphases of the presentation are threefold: (1) a brief articulation of some conceptual differences between design-based and model-based approaches, (2) examples that illustrate these differences, and (3) comments on particular advantages and disadvantages of both approaches.

49 Missing Information in Health-Related Quality-of-Life Data ●

Section on Health Policy Statistics, ENAR, Biometrics Section, WNAR

Sunday, July 29, 4:00 pm–5:50 pm

Bivariate Normal Frailty Models for Multiple QOL Outcomes

❖Andrea B. Troxel, University of Pennsylvania, 423 Guardian Drive, School of Medicine, 632 Blocklwy Hall, Philadelphia, PA 19104, *atroxel@cceb.med.upenn.edu*; Denise A. Esserman, The University of North Carolina at Chapel Hill

Key Words: frailty models, quality of life, multivariate survival, grouped data

This talk describes bivariate normal frailty models, with application to multiple, grouped, quality-of-life (QOL) events observed in clinical trials. Multiple QOL scales are assessed, with interval censoring, or grouping, due to the QOL assessment schedule. A bivariate normal frailty is introduced that incorporates dependence among the events while allowing different types of events to have different frailties (e.g., emotional vs. physical functioning scales). Extensions to joint analysis of QOL and survival outcomes will be discussed. Simulation results and a QOL example will be presented.

The Analysis of Questionnaire Submission Rates in Longitudinal HRQOL Assessment

❖Stephanie R. Land, University of Pittsburgh, 201 N. Craig Street, Suite 350, Pittsburgh, PA 15213, *land@pitt.edu*; Marcie W. Ritter, University of Pittsburgh; Sarah Haile, University of Pittsburgh

Key Words: quality of life, patient-reported outcomes, mixed effects model, compliance, cancer, clinical trials

In longitudinal clinical studies, tremendous resources are often required to assure continued compliance with study assessments. This presentation is motivated by examples of HRQOL studies in the setting of multi-center cancer clinical trials. In these trials, participants are asked to complete questionnaires at fixed intervals over several years. We seek a greater understanding of both the uncontrolled factors that may influence compliance (e.g. participant age), and the factors we can control, such as clinic newsletters. I will present methodology and results from a randomized trial (nested within a cancer clinical trial) we have undertaken to evaluate the effectiveness of a compliance intervention. The analysis uses a mixed model approach, in which the submission of a questionnaire is the binary outcome variable. I will also discuss design considerations for compliance studies.

The Value of Ancillary Data in Longitudinal Studies of Health-Related Quality of Life with Informative Dropout

❖Diane Fairclough, University of Colorado at Denver and Health Services Center, Mail Stop F443, PO Box 6508, Aurora, CO 80045-0508, *diane.fairclough@uchsc.edu*; Mark Jaros, University of Colorado at Denver and Health Services Center

Key Words: Missing data, longitudinal studies, Shared parameter models, Informative Dropout, Quality of LIfe

Assessment of change in populations with morbidity and mortality is complicated by dropout by individuals experiencing negative effects of disease and side effects of treatment. Analytic strategies include mixture models that condition on the pattern of missing assessments, the time to dropout or an event associated with dropout. These later models jointly estimate the trajectory of the longitudinal outcome and the time to the event linking the models through shared parameters or random effects. Early applications utilized survival and time to dropout as the ancillary data. The aim of this paper is to illustrate how other ancillary data, proximal to the outcome of interest in the causal pathway, may improve estimation of change. Careful study planning including collection of ancillary data will allow joint models which are conditionally MAR and reduce bias associated with dropout.

50 Issues in Conducting Experiments in Statistical Education

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Sunday, July 29, 4:00 pm–5:50 pm

Issues in Conducting Experiments in Statistical Education

❖ William Harkness, The Pennsylvania State University, 318 Thomas Bldg, University Park, PA 16802, *wlh@stat.psu.edu*; ❖ G. Bruce Schaalje, Brigham Young University, 230 TMCB, Provo, UT 84602, *schaalje@byu.edu*; ❖ Dennis K. Pearl, The Ohio State University, 404 Cockins Hall, 1958 Neil Ave, Columbus, OH 43210-1247, *pearl.1@osu.edu*; ❖ John P. Holcomb, Cleveland State University, Department of Mathematics, 2121 Euclid Ave., RT 1512, Cleveland, OH 44115-2214, *j.p.holcomb@csuohio.edu*

Key Words: Education, Stat Ed Research, Experiments, Assessment

GAISE gives recommendations for improving statistical education, but does implementing these guidelines improve statistical education? How can we test whether changes do indeed improve students' knowledge and understanding? How can a clinical trial be performed? Panelists will discuss challenges that they faced in conducting experiments in statistical education and give their recommendations for others wishing to conduct such research.

51 Ethnic Comparisons in Global Health Research ● ✪

Biopharmaceutical Section, Section on Health Policy Statistics

Sunday, July 29, 4:00 pm–5:50 pm

Assessment of Similarity in Dose Responses Between Two Ethnic Populations for Drug Development: A Resampling Approach

❖ Kyoungah See, Eli Lilly and Company, 3248 Blue Ash Lane, Indianapolis, IN 46239, *seeky@lilly.com*

Key Words: Analysis of covariance, Bone mineral density, Bridging, Osteoporosis, Power Curves

According to ICH, data from an existing study may be extrapolated to a new study, in a new region utilizing a different dose, if the safety and efficacy profiles are not substantially different. One of the main challenges of this type of study, based on foreign clinical trials, is the difference in sample sizes. Specifically, the Global data, which is frequently the original data in such studies, regularly contains a much larger sample size than its Asian (or local) counterpart. In order to study the similarity of a drug response in the two ethnic populations, we investigate resampling techniques to overcome the size disparity between the data sets, while utilizing a general linear model to adjust the fundamental differences between the two populations, with respect to some important covariates. We provide an example of their utility based on clinical trials of raloxifene.

The Application of Adaptive Design in R&D in Japan from a Statistical Point of View

❖ Hajime Uno, Kitasato University, 5-9-1, Shirokane, Minato-ku, Tokyo, 108-8641 Japan, *unoh@pharm.kitasato-u.ac.jp*; Masahiro Takeuchi, Kitasato University; Fumiaki Takahashi, Kitasato University

Key Words: global trial, bridging study

The U.S. Food and Drug Administration published "Challenge and Opportunity on the Critical Path to New Medical Products" in 2004 in order to stimulate an innovative approach to the research and development of new medical products. It is well recognized that the excellent progress of basic science has discovered potential products, but that the slow progress of applied science has failed to deliver them to patients in a timely fashion. Therefore the concept of an adaptive design has been investigated extensively. Recently, the concept of the E5 guideline has been extended to a simultaneous global study through mutual extrapolation of clinical data among regions. We investigate the application of adaptive design in R&D in Japan and discuss the associated statistical issues related to usage of surrogate markers, sample size re-calculation and the definition of similarity of drug efficacy.

A Bayesian Approach to Evaluation of Bridging Evidence

❖ Hsiao-Hui Tsou, National Health Research Institutes, No. 35, Keyan Road, Zhunan Town,, Miaoli County, Zhunan, 35053 Taiwan, *tsouhh@nhri.org.tw*; Chin-Fu Hsiao, National Health Research Institutes; Yu-Yi Hsu, Iowa State University; Jen-pei Liu, National Taiwan University

Key Words: Bridging Study, Bayesian Approach, Extrapolation, Similarity

In this talk, we give some successful examples of bridging evaluation conducted in Taiwan. For evaluations of bridging studies, Liu, Hsiao and Hsueh (2002) have proposed a Bayesian approach to synthesize the data generated by bridging study and foreign clinical data generated in the original region for assessment of similarity based on superior efficacy of the test product over a placebo control. However, the results of the bridging studies using their method are overwhelmingly dominated by the results of the original region due to an imbalance of sample sizes between the regions. Therefore, we propose a Bayesian approach using a mixture prior for assessment of similarity between the new and original region based on the concept of positive treatment effect. Methods for sample size determination are proposed. Numerical examples illustrating uses of the proposed procedures are provided.

Minority Differences in Cancer Survival on Cooperative Group Clinical Trials

❖ Beow Yong Yeap, Massachusetts General Hospital/Harvard Medical School, 50 Staniford St, Suite 560, Boston, MA 02114, *byeap@partners.org*; Marvin Zelen, Harvard School of Public Health

Key Words: African-American, Cancer clinical trials, Survival, matching

Outcome comparisons between racial groups generally lack the individual data to adjust for patient variation in disease, therapy and institution factors. The Eastern Cooperative Oncology Group (ECOG) has conducted multi-center clinical trials in adult cancers nationwide since 1955. We use the ECOG database to determine the outcome of African-American (AA) patients on therapeutic protocols during 1976-2005. To minimize patient variation without explicit covariate modeling, each AA proband was matched by gender to a White patient enrolled on the same protocol, assigned to the same arm and treated at the same hospital. We previously found AA patients experienced a higher death rate, with the deficit distributed disproportionately among AA women. Using a larger sample size and extended follow-up, we examine whether the excess mortality of AA patients has diminished on ECOG clinical trials.

52 Ranked Set Sampling II ✪

Section on Nonparametric Statistics, Section on Survey Research Methods

Sunday, July 29, 4:00 pm–5:50 pm

Optimal Sign Test for One Sample Bivariate Location Model Using an Alternative Bivariate Ranked Set Sample

❖Hani Samawi, Georgia Southern University, Department of Mathematical Sciences, POBox 8093, Statesboro, CA 30640, *hsamawi@georgiasouthern.edu*

Key Words: Bivariate Ranked Set Sample, Location Model, Median Ranked Set Sample, Pitman efficiencies, Sign Test

The aim of this paper is to find optimal alternatives bivariate ranked set sample for one sample location model bivariate sign test. Our numerical and theoretical results indicated that the optimal designs for bivariate sign test are the alternative designs with quantifying order statistics with labels [(r+1)/2, (r+1)/2] when the set size r is odd and {[(r/2)+1, r/2], [r/2, (r/2)+1]} when the set size r is even. The asymptotic distribution and Pitman efficiencies of those designs are derived. Simulation study is conducted to investigate the power of the proposed optimal designs. Illustration using real data with Bootstrap algorithm for P-value estimation is used.

Optimal Allocation for Unbalanced Stratified Ranked Set Sample

❖Chris Sroka, The Ohio State University, 1958 Neil Avenue, Cockins Hall, Room 404, Columbus, OH 43210-1247, *csroka@stat.ohio-state.edu*; Elizabeth A. Stasny, The Ohio State University; Douglas Wolfe, The Ohio State University

Key Words: Neyman allocation, stratification, judgment order statistic, variance estimation

Stratified ranked set sampling is a data collection method that stratifies the population before conducting ranked set sampling (RSS). It is similar to traditional stratified sampling, except that RSS is used in place of simple random sampling. This method of sampling results in more precise estimators of the mean than traditional stratified sampling or RSS alone. This sampling method raises an interesting question about how to allocate a fixed sample size to both the strata and the ranks within each stratum. We examine possible allocation schemes to find the method that produces the lowest variance of the mean estimator.

Order-Restricted Randomized Designs for Linear Models Using L1 Norm

❖Shannon Markiewicz, The Ohio State University, 955 Mount Pleasant Ave, Columbus, OH 43201, *markiewicz.8@osu.edu*; Omer Ozturk, The Ohio State University

Key Words: order restricted randomized design, ranked set sampling, L1 Norm

Ranked set sampling is not well-suited for the design of experiments primarily for two reasons; it requires larger number of experimental units, and the role of randomization is not well defined. To resolve these concerns, recently order restricted randomized (ORR) designs have been developed and its properties have been discussed in the literature. In this paper, we develop nonparametric statistical inference for linear models in the context of order restricted randomized designs based on L1 norm. We derive the asymptotic distribution of the parameter estimates, and develop drop, Wald, and score tests for the contrast parameter. It is shown that ORR design performs better than its competitor in the literature. The Type I error is examined along with power comparisons. The properties of the estimators and tests are discussed under an arbitrary but consistent ranking scheme.

Optimal Allocation in the Presence of Missing Data

❖Jessica Kohlschmidt, The Ohio State University, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *jessica@stat.ohio-state.edu*; Elizabeth A. Stasny, The Ohio State University; Douglas Wolfe, The Ohio State University

Key Words: Ranked Set Sampling, Missing data, Optimal Allocation, MCAR, MAR

Missing data is a natural consequence of sampling. Some researchers ignore the missing data. If the data is missing completely at random, this procedure provides meaningful estimates. In many situations, the missingness in the data is due to a phenomenon occurring in the population. In ranked set sampling, we stratify the data by ranks. Many times the missingness in the population varies depending on which rank the observation came from. We will show the optimal allocation when the missing data parameters are known. Then we extend this to the case where the missing data parameters are unknown, a more convincing scenario. We explore the effect of the varying costs for collecting observations in each rank and its effect on the optimal allocation.

Rejective Sampling and Conditional Tests from Biased Samples for the Equality of Two Distributions

❖Qing Kang, North Dakota State University, Waldron E1A, Dept of Statistics NDSU, Fargo, ND 58105, *qing.kang@ndsu.edu*; Paul I. Nelson, Kansas State University

Key Words: weighted distributions, nonparametric test, asymptotic normality, stochastic ordering, location shift

This study proposes a class of nonparametric tests for the equality of two distributions when samples are collected independently under different, biased sampling schemes. Asymptotic properties are derived using a central limit theorem developed for rejective samples from a finite population. Conditions under which the tests are consistent are given.

53 Consumer Product Risk ● ✪

Section on Risk Analysis

Sunday, July 29, 4:00 pm–5:50 pm

Sex Differences: Rare Traffic Fatalities Support Nature over Nurture

❖Leonard Evans, Science Serving Society, 973 Satterlee Road, Bloomfield Hills, MI 48304, *le@scienceservingsociety.com*

Key Words: sex differences, gender differences, traffic fatalities, nature versus nurture, child traffic fatalities, baby traffic deaths

As females and males are treated differently in all societies from birth, socialization explanations can always be offered for any gender dependent outcomes. Traffic fatality data are used to explore gender-dependent outcomes that do not appear to have plausible socialization explanations. Rare events are extracted from 1.3 million traffic fatalities from 1975 through 2003. It is found that 17 children aged 5 or under were killed as sole occupants sitting in driver seats of vehicles traveling on public roads. The gender mix was 17 boys, no girls. More male than female pedestrians were killed at all ages, including the first year of life. Similar gender and age dependence is found for testosterone measured in saliva, arrest rate, and traffic-crash rate. The results are interpreted to support innate behavioral differences originating at a hormonal level.

Evaluation of the Effect of ABS in Truck Crashes Using LTCCS Data

❖ Wai Chan, Exponent, Inc., 21 Strathmore Road, Natick, MA 01760, *wchan@exponent.com*; Rose Ray, Exponent, Inc.; Bruce Ketcham, Exponent, Inc.

The Large Truck Crash Causation Study (LTCCS) is based on a nationally representative sample of 1132 large trucks involved in crashes with fatal or serious injury. The study period is the 33 months between 04/01 and 12/03. It provides a comprehensive description of the physical events, along with the vehicles and drivers information. This paper examines the effect of antilock brake systems (ABS) on trucks involved in single and multi-vehicle crashes. Federal Motor Vehicle Safety Standards were amended by the NNHTSA in 96 to require the installation of ABS on new medium and heavy trucks, truck-tractors and trailers. ABS is intended to help maintain directional stability and control during braking. We explored the crashes from the LTCCS data that involves loss of control, skidding and jackknife and compared the crash rates for ABS and non-ABS vehicles.

National Estimates of Nonfatal Consumer Product Injuries: Surveillance Data from Hospital Emergency Departments, 1980–2005

❖ Madhu Iyer, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *miyer@exponent.com*

Key Words: Survey Design, Data Collection, Sampling Error, Non-Sampling Error

The primary source of data in the US on product-associated injuries is the National Electronic Injury Surveillance System (NEISS) maintained by the Consumer Product Safety Commission (CPSC). NEISS monitors emergency room visits associated with consumer products. The monitoring is carried out using a small (about 2%) but statistically representative set of hospital emergency rooms throughout the United States and its territories. This system provides the CPSC with the majority of its injury data and is used to estimate the total number of product-related injuries treated in all hospital emergency rooms in the US. This talk will provide the general structure of the NEISS sampling design and its strategy for data collection, discuss sampling and nonsampling errors, and touch on examples that demonstrate the usage of the database.

To Recall or Not To Recall? Statistical Assessment of Consumer Product Failure Risk

❖ Ke Zhao, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *kzhao@exponent.com*; Duane Steffey, Exponent, Inc.

Key Words: Product, Recall, Failure, Risk, Manufacturer

Unintended declines in the production quality of consumer items can pose safety hazards to consumers, increase failure rates, and shorten product life - which, in turn, causes unexpected financial burdens and potential legal liabilities to manufacturers. In monitoring the field performance of its products, a manufacturer needs to determine whether an adverse change has occurred and, if so, assess the implications for product safety and reliability. A decision to recall sold product must involve identification of the units affected, typically by time or source of production. Drawing on experience with diverse products and industries, we illustrate how statisticians can play an important role in such investigations by framing the issues in a conceptual model, extracting information from imperfect manufacturing data, and communicating findings to upper management and regulatory agencies.

54 Analyzing Time Series Data with Time-Varying Frequencies ●

Section on Physical and Engineering Sciences, Section on Statistics and the Environment

Sunday, July 29, 4:00 pm–5:50 pm

The Application of the Kalman Filter to Nonstationary Time Series Through Time Deformation

❖ Zhu Wang, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, LE-400, Seattle, WA 98109-1024, *zwang@scharp.org*

Key Words: nonstationary time series, time deformation, kalman filter

An increasing valuable tool for modeling nonstationary time series data is time deformation. However, since the time transformation transforms the scale, equally spaced data become unequally spaced data. Interpolation is therefore often used to obtain regularly sampled data, which can be modeled by the classical ARMA modeling techniques. In this paper, the need for interpolation is eliminated by employing the continuous time autoregressive model. To estimate the parameters, the maximum likelihood function is decomposed by means of the Kalman filter. The resulting modeling improvements include, more accurate estimation of the spectrum, better forecasts, and the separation of the data into its smoothed time-varying components. The technique is applied to simulated and real data for illustrations.

Time-Frequency Analysis: G(lambda)-stationary Processes

❖ Huiping Jiang, Columbia University, 1051 Riverside Drive, Unit 48, New York, NY 10032, *hj2179@columbia.edu*; Henry L. Gray, Southern Methodist University; Wayne Woodward, Southern Methodist University

Key Words: time-frequency, M-stationary process

Methods such as wavelets and M-stationary process have been developed to analyze the time-frequency properties of a process where frequency changes with time. In certain circumstances, when the frequencies of a process change systematically either monotonically increasing or monotonically decreasing across time, another approach is to apply an appropriate Box-Cox transformation to the time axis for the given signal in order to obtain a new stationary data set. This new data set can be analyzed by standard methods. Processes which are transformed to a stationary process after Box-Cox transformation on the time scale are called G(lambda)-stationary processes, where lambda is the corresponding parameter of the Box-Cox transformation.

Filtering Analysis for Nonstationary Time Series by Time Deformation

❖ Mengyuan Xu, Southern Methodist University, 6061 Village bend Dr, Apt 514, Dallas, TX 75206, *txu@mail.smu.edu*

Key Words: time-deformation, non-stationary, filter, fourier transform, linear chirp, time series

Classical linear filtering theories and methodologies can only be used in stationary process where frequencies do not vary with time. In this paper, we will discuss for those non-stationary processes, in which the frequencies change with time in a certain way, such as G(?) and linear chirp processes, how to do filtering by applying time deformation. Then, we will compare this filtering method with other filtering methods for non-stationary time series such as fractional Fourier transform method, which is specially used for linear chirp signals. In this paper, we will also provide a solution to end-effect problem for linear filter.

Generalization and Applications of Linear Chirp Stationary Processes

❖Stephen Robertson, Southern Methodist University, 1600 Abrams, Number 18, Dallas, TX 75214, *sdrobert@smu.edu*

Key Words: G-Stationary Processes (GSP), Linear Chirp Stationary Processes (LCSP)

Time deformation has been proposed as an alternate method to window-based approaches when analyzing time series data with time-varying frequency. Specifically, "G-Stationary Processes" (GSP) are processes which are transformed to a stationary process after a transformation of the time scale, thereby allowing use of standard approaches for analyzing stationary data. A special case of a GSP is the Linear Chirp Stationary Process (Liu, Gray, Woodward, 2004), which is a non-stationary process whose frequency changes approximately linearly in time. We generalize previous assumptions to allow analysis of LCSP for a wider range of applications. The LCSP can be classified into "generalized" and "LC(p,q)" chirp processes. We show that the relationship between generalized and LC(p,q) chirp processes is analogous to that of harmonic and ARMA models. Results are applied to simulated and actual data.

Time Transformation Methods for Analyzing Data with Time-Varying Frequencies, with Applications

❖Wayne Woodward, Southern Methodist University, Department of Statistical Science, Dallas, TX 75275, *waynew@smu.edu*; Henry L. Gray, Southern Methodist University

Key Words: time deformation, TVF data, nonstationary, time series

Many time series data sets are characterized by frequency behavior that changes with time. Examples of this type of data are Doppler signals (for measuring blood flow, etc.), heart sounds, bat echolocation signals, seismic waveforms, etc. Time series of this type are nonstationary, and the standard methods based on a stationarity assumption are not appropriate. Techniques that have been used for analyzing time varying frequency (TVF) data include wavelets and short term Fourier transforms. When the frequencies of a process are changing continuously in time, these methods may not apply. In this case a new approach based on the transformation of time may be appropriate. The strategy is to transform the time axis in order to obtain a new stationary realization that can then be analyzed by standard methods (spectral analysis, filtering, etc.).

55 Spatial Statistics: Methods and Applications ✪

IMS, Section on Statistical Graphics, Section on Statistics and the Environment

Sunday, July 29, 4:00 pm–5:50 pm

Exploratory Spatial Data Analysis Using APLE Statistics

❖Hongfei Li, The Ohio State University, 1958 Neil Ave, Cockins Hall, Columbus, OH 43210, *hongfei@stat.ohio-state.edu*; Catherine A. Calder, The Ohio State University; Noel Cressie, The Ohio State University

Key Words: Spatial dependence parameter, SAR model, Profile likelihood, Exploratory spatial data analysis

In exploratory analyses of spatial data, typically the primary goal is to asses whether the data exhibit spatial dependence and then, if so, to determine the strength of the spatial dependence. In this talk, we consider a class of statistics called APLEs, which are approximate profile likelihood estimators in common spatial statistical models. Importantly, APLE statistics are constructed to have a closed form. We provide both theoretical and simulation-based evidence to demonstrate that APLEs perform well as both estimators and test. In addition, for the APLE statistic based on SAR model, we propose graphical exploratory tools, which allow us to assess the strength of spatial dependence in the data as a whole and to identify areas of local spatial clustering and outliers. Finally, we use real data to illustrate the use of APLE as an exploratory spatial-data-analysis tool.

Statistical Trend Detection of a Global Change Signal in Regional Climate

❖Airong Cai, University of Illinois at Chicago, IDS Department MC294 601 S Morgan St, Chicago, IL 60616, *acai1@uic.edu*; Katharine Hayhoe, Texas Tech University; George Tiao, The University of Chicago; Donald Wuebbles, University of Illinois at Urbana-Champaign

Predicting near-term (5- to 30-year) trends in regional climate is inherently difficult. How long will it take before the upward trend in temperature becomes significant at the regional level? Over what time horizon will decision makers from various regions need to take into account climate change projections for that region? To address these questions we extend the technique of statistical trend detection proposed by Tiao et al. (1990) by generalizing the underlying annual trend statistical model to accommodate seasonal trends with seasonal variances that are more proper for temperature data analysis. The extended technique is then applied to observed historical data and model-simulated temperature data for grid cells across the globe as well as US.

Detection of the Significance of Cool Spot Clusters with Application to the Indiana Colorectal Cancer Study

❖Tonglin Zhang, Purdue University, 250 N University Street, West Lafayette, IN 47906-2066, *tlzhang@stat.purdue.edu*

Key Words: Clustering and Cluster, Cool spot, Hot Spot, Permutation Test

The detection of spatial clustering and clusters for disease incidence and mortality is an essential problem in spatial epidemiology and statistical approaches have been of great interest over the past 20 years. However in the literature, most research articles focuses on the problems of the detection of hot spot clusters and ignore the problems of the detection of cool spot clusters. In this paper, we propose a method of the detection of cool spot clusters. In our case study of the Indiana colorectal cancer study, we successfully detected a cool spot around the Indianapolis area and the statistical analysis showed that this cluster was the only local cluster in the dataset.

Spatial Statistical Analysis of Doctors' Prescription Amounts by Region

❖Lei Kang, The Ohio State University, Department of Statistics, 1958 Niel Avenue 404 Cockins Hall, Columbus, OH 43210-1247, *lei@stat.ohio-state.edu*; Noel Cressie, The Ohio State University; Desheng Liu, The Ohio State University

Key Words: EDA, ESDA, CAR model, Maximum likelihood estimation, Spatial dependence

In this talk, we analyze doctors' prescribing patterns based on regional aggregation of individual prescription data. Through a series of exploratory-data-analysis (EDA) techniques, we demonstrate the importance of transforming the data and of doing a weighted regression analysis. Both a non-spatial model and a conditional autoregressive (CAR) model are fitted. Our results show that CAR models have smaller sum of squared prediction errors compared to non-spatial models. Different decompositions of large- and small-scale variation are considered; we see that when a lot of spatial variation is attributed to the large-scale, deterministic component, spatial-dependence parameters describing the small-scale, stochastic component decrease in value with larger standard errors. This work motivates the use of a hierarchical spatial statistical model to describe doctors' prescribing patterns.

Mis-specified Maximum Likelihood Estimates and Spatial Error Models

❖ Jiehua Chen, Stanford University, Statistics Department, Stanford, CA 94305, *chenjh@stanford.edu*; Paul Switzer, Stanford University

The paper investigates the impact of covariance model misspecification on the estimates of the regression coefficients in a spatial error model. Spatial error models are the regression models which have spatially indexed observations and spatially correlated residuals. Specifically, we evaluate the asymptotic bias for the regression coefficients and estimates of the standard errors of coefficients that results from incorrect assumptions regarding the covariance matrix of the error term. The asymptotic bias is the lower bound of the bias and the paper also tries to shed some lights on the possible upper bound of bias by simulations.

56 Some Recent Advances in High-Dimensional Data Modeling ✪

IMS, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, WNAR

Sunday, July 29, 4:00 pm–5:50 pm

Local Multidimensional Scaling for Nonlinear Dimension Reduction, Graph Layout, and Approximity Analysis

❖ Lisha Chen, Yale University, 24 Hillhouse Ave, New Have, 06511, *lisha.chen@yale.edu*; Andreas Buja, University of Pennsylvania

Key Words: nonlinear dimension reduction, MDS, graph layout, proximity analysis

Dimensional reduction has many applications in science. Classical methods like PCA or MDS fail to recover the intrinsic dimensions of the data if there is nonlinear structure in them. There has been a resurgence of interest in nonlinear dimension reduction (NLDR) methods. We will review some of the recently proposed methods and introduce a family of new nonlinear dimension reduction methods called "Local Multidimensional Scaling" or LMDS. LMDS only uses local information from user-chosen neighborhoods like other NLDR methods, but it differs from them in that it uses the force paradigm from graph layout by proposing a parameterized family of stress or energy functions. This family provides users with considerable flexibility for achieving desirable embeddings. Facing an embarrassment of riches of energy functions, we propose a metacriterion for selecting viable energy functions.

Model-Based Protein Summaries and Differential Label-Free Quantitative Proteomics

❖ Alan Dabney, Texas A&M University, 3143, College Station, TX 77845, *adabney@stat.tamu.edu*

Key Words: proteomics, missing data, censored regression, differential expression

Differential label-free quantitative proteomics refers to the use of MS peak areas to quantitate and compare protein expression levels in LC-MS/MS experiments. There is typically widespread informative missingness due to detection limits and other factors. I propose a method for rolling peptide information up to the protein level while accounting for the information in missing peaks. Relationships across peptides are exploited to facilitate estimation and stabilize estimates. Real examples will be illustrated with data from the Pacific Northwest National Laboratory.

Kernel-Induced Classification Tree and Random Forest

❖ Guangzhe Fan, University of Waterloo, 200 University Ave W, Department of Statistics, Waterloo, ON N2L 3G1 Canada, *gfan@uwaterloo.ca*

Key Words: classification, kernel, tree, random forest, learning

A recursive-partitioning procedure using kernel functions is proposed for classification problems. We call it KICT: kernel-induced classification trees. The resulting model could significantly improve the traditional CART model in many situations. We also introduce KIRF: kernel-induced random forest. KIRF also compares favorably to the traditional random forests in many situations. We use simulated and real world data to illustrate their performances.

A Statistical Perspective on Nonlinear Dimensionality Reduction and Manifold Learning

❖ Ann Lee, Carnegie Mellon University, Dept of Statistics BH 229J, Pittsburgh, PA 15213, *annlee@stat.cmu.edu*

Key Words: manifold learning, eigenmaps, kernel PCA

In some problems, the dimension d of the given input space is very large while the data itself has a low intrinsic dimensionality. There are several so called manifold learning methods that aim to learn the geometry of nonlinear structures embedded in R^d. Most of these methods, however, approach the data as a fixed set of points and use deterministic techniques to find the optimal embedding coordinates. In this talk, I will present an alternative approach to manifold learning as a statistical estimation problem. Issues such as random noise and regularization will be discussed and put into the context of eigenmaps and kernel PCA.

Some Recent Advances in High-Dimensional Data Modeling

❖ Wenbin Lu, North Carolina State University, 2501 Founders Drive, 210E Patterson Hall, Raleigh, NC 27695-8203, *lu@stat.ncsu.edu*; Hao Zhang, North Carolina State University

Key Words: Adaptive lasso, Penalized likelihood, Proportional hazards model, Proportional odds model, Survival data, Variable selection

We study the problem of variable selection for a class of semiparametric survival models with censored data. A unified model selection and estimation procedure with desired theoretical properties and computational convenience is proposed. Numerical results indicate that the new method can produce accurate and interpretable models, and performs competitively.

57 Statistical Machine Learning and High-Dimensional Inference ✪

IMS, Section on Nonparametric Statistics

Sunday, July 29, 4:00 pm–5:50 pm

Nonparametric Density Estimation in High Dimensions Using the Rodeo

❖ Han Liu, Carnegie Mellon University, Baker Hall 132, Pittsburgh, PA 15213, *hanliu@stat.cmu.edu*; John Lafferty, Carnegie Mellon University; Larry Wasserman, Carnegie Mellon University

Key Words: nonparametric inference, density estimation, sparsity, high dimensionality, bandwidth selection, rodeo

We consider the problem of estimating the joint density of a d-dimensional random vector $X = (X_1,...,X_d)$ when d is large. We assume that the density is a product of a parametric component and a nonparametric component which depends on an unknown subset of the variables. Using a modification of a recently developed nonparametric regression framework called rodeo (regularization of derivative expectation operator), we propose a method to greedily select bandwidths in a kernel density estimate. It is shown empirically that the density rodeo works well even for very high-dimensional problems. When the unknown density function satisfies a suitably defined sparsity condition, and the parametric baseline density is smooth, the approach is shown to achieve near optimal minimax rates of convergence, and thus avoids the curse of dimensionality.

Functional Data Analysis in Lipoprotein Profile Data

❖ Yehua Li, University of Georgia, 204 Statistics Building, 101 Cedar Street, Athens, GA 30602, *yehuali@uga.edu*; Tailen Hsing, The Ohio State University

Key Words: Functional Data Analysis, linear model, Kalman filter, principal component, projection pursuit, penalized spline

Lipoprotein concentration in human serum is an important risk factor for cardiovascular heart disease. Different species of lipoprotein in a serum sample can be separated by a centrifugation treatment, according to their density differences. A lipoprotein profile for a patient is obtained by taking the image of his centrifuged serum sample, after an application of a lipophilic stain. In this paper, we use functional data analysis techniques to study the lipoprotein profile curves. The goal is to predict the quantity of total cholesterol and each species of lipoprotein from the profile curves. We discuss implementation issues including smoothing the profile curves with P-spline using the Kalman filter technique, functional PCA and functional linear model. We also study a functional projection pursuit model, as a device to check the adequacy of the functional linear model.

Variable Selection Procedures for Generalized Linear Mixed Models in Longitudinal Data Analysis

❖ Hongmei Yang, North Carolina State University, 3002 Kings Ct Apartment C, Raleigh, NC 27606, *hyang3@ncsu.edu*; Daowen Zhang, North Carolina State University; Hao Zhang, North Carolina State University

Key Words: Smoothly clipped absolute deviation, Penalized quasi-likelihood, Longitudinal data, Restricted maximum likelihood, Generalized linear mixed models, Variance components

For non-sparse longitudinal data such as count data and binomial data with moderate to large binomial denominators, we propose Penalized Quasi-Likelihood (PQL) procedure for simultaneous model selection and estimation. Due to the low estimation ability of PQL for binary data, we propose three other procedures: Full Likelihood Model Selection (FLMS), Two-stage Penalized Quasi-Likelihood Model Selection (TPQLMS) and approximate Marginal Likelihood Model Selection (AMLMS). Among them, FLMS and PQLMS have the feature of selecting informative variables and estimating regression parameters simultaneously. A robust estimator of standard deviation is derived based on a sandwich formula and tested through simulations for FLMS and PQLMS. A bias correction is proposed to improve the estimation accuracy of PQLMS.

Semiparametric Estimation of Covariance Matrices for Longitudinal Data

❖ Yichao Wu, Princeton University, Dept of ORFE, Princeton, NJ 08544, *yichaowu@princeton.edu*; Jianqing Fan, Princeton University

Key Words: Correlation structure, difference-based estimation, longitudinal data, quasi-maximum likelihood, varying-coefficient partially linear model

Estimation of the covariance structure for longitudinal data poses significant challenges, as they are usually collected at irregular time points. Fan, Huang, and Li (07) proposed using a varying-coefficient partially linear model coupled with a semiparametric cov structure. After estimating other components, they estimated the parametric correlation structure via quasi-maximum likelihood method and showed its good finite-sample performance based on extensive simulation studies. In this work, we study the semiparametric varying-coefficient partially linear model under weaker smoothness conditions on the varying regression coefficient function, propose to estimate regression coefficients using a newly developed difference-based technique, and provide large-sample properties of our estimators. The particular focus is given to asymptotic properties of the quasi-maximum likelihood estimator.

Characterizing the Function Space for Bayesian Kernel Models

❖ Natesh Pillai, Duke University, 1022 Sedgefield Street, Durham, NC 27705, *nsp2@stat.duke.edu*

Key Words: RKHS, non-parametric Bayes, Levy processes, Integral operators

Kernel methods have been very popular in the machine learning literature in the last ten years, often in the context of Tikhonov regularization algorithms. In this paper we study a coherent Bayesian kernel model based on an integral operator whose domain is a space of signed measures. Priors on the signed measures induce prior distributions on their image functions under the integral operator. We study several classes of signed measures and their images, and identify general classes of measures whose images are dense in the reproducing kernel Hilbert space (RKHS) induced by the kernel. This gives a function-theoretic foundation for some nonparametric prior specifications commonly-used in Bayesian modeling, including Gaussian processes and Dirichlet processes, and suggests generalizations. We outline a general framework for the construction of priors on measures using LÈvy processes.

58 Bayesian Hierarchical Modeling and Computation ● ✪

Section on Bayesian Statistical Science, WNAR
Sunday, July 29, 4:00 pm–5:50 pm

Latent Variable Model for Multiple Outcomes with Nonignorable Missing Data

❖ Xiaohong Yan, University of California, Los Angeles, Department of Biostatistics, School of Public Health, Los Angeles, CA 90095, *xhyan@ucla.edu*; W. John Boscardin, University of California, Los Angeles

Key Words: latent variable, multiple outcomes, non-ignorable missing data, multivariate longitudinal data

In clinical trials, multiple outcomes are usually collected over time to measure the same quantity of interest, such as treatment effect. These measures could be in different data type (discrete or continuous), and tend to be incompletely recorded due to drop out, death or other reasons. Ignoring these missing data might induce bias in inference analysis. We propose a latent variable shared parameter model for multiple outcomes in different data type (binary and continuous) with non-ignorable missing data. The latent variable accommodates the correlation structure among different response outcomes. The mechanism for those non-ignorable missing data is modeled by a shared parameter model. We fit the model using Markov chain Monte Carlo method. Our method is illustrated using real data from scleroderma lung study (SLS).

Modeling Common Effects of Predictors on Multiple Longitudinal Outcomes

❖ Juan Jia, University of California, Los Angeles, Department of Biostatistics, School of Public Health, Los Angeles, CA 90025, *jiajuan@ucla.edu*; Robert E. Weiss, University of California, Los Angeles

Key Words: Multivariate longitudinal data, Bayesian inference, common effect

Longitudinal data with multivariate outcomes measured over time are common in medical, psychological and sociological fields. In typical models, each covariate has a different effect on each outcome. However, outcomes are often quite similar and covariate effects might be expected to be similar as well. Lin et al. (2000) estimated a common effect for single covariate. We instead propose a model to evaluate a common effect for the entire linear predictor. Maximum a posteriori inference with a flat prior is used to estimate model parameters. We apply the proposed method to the Brief Symptom Inventory (BSI) data from children of HIV+ parents. We estimate common effects of age, gender, and parental drug use on nine sub-scales.

Using Incompatible Gibbs Samplers To Efficiently Fit Hierarchical Models

❖ Taeyoung Park, University of Pittsburgh, 2717 Cathedral of Learning, 4200 Fifth Avenue, Pittsburgh, PA 15260, *tpark@pitt.edu*; David van Dyk, University of California, Irvine

Key Words: Bayesian data analysis, Gibbs sampler, Hierarchical models, Incompatibility, MCMC

The Gibbs sampler is a very useful technique for fitting highly structured hierarchical models that are now prevalent in biological, physical, engineering, and social sciences. Its sometimes slow convergence, however, has been a longstanding complaint especially when complex models are fitted. To alleviate the slow-mixing behavior, we present efficient Gibbs sampling strategies by capitalizing on incompatible conditional distributions derived from different marginal distributions of a target distribution. Until now, such incompatibility has been simply avoided in the construction of Gibbs samplers because the resulting convergence properties were not well understood. However, incompatible Gibbs samplers have known stationary distributions as well as fast convergence. Their applications are illustrated using a highly structured multilevel spectral model in astrophysics.

Bayesian Mixtures for Modeling Correlation

❖ Lei Qian, University of California, Los Angeles, 3110 Sawtelle Blvd, #308, Los Angeles, CA 90066, *leiqian@ucla.edu*; Robert E. Weiss, University of California, Los Angeles

Key Words: Covariance model, Longitudinal data, Correlation model, MCMC

In longitudinal data analysis, correctly modeling the covariance matrix is important. Simple structured covariance models have advantages in parameter dimensionality and inference, but it is not necessarily right for complex stochastic processes. We propose a Bayesian correlation mixture model which provides a rich family of correlation models. We decompose the covariance matrix into variances and correlations components. The correlation matrix is modeled by a convex combination of two or more simple correlation matrices. Examples and simulation studies illustrate how our model works with hypothetical and real data correlations.

59 Section on Bayesian Statistics Student Paper Competition: New Developments in Bioinformatics and Genetics ● ✪

Section on Bayesian Statistical Science

Sunday, July 29, 4:00 pm–5:50 pm

Hidden Markov Model for Jointly Modeling Probe Sequences and ChIP-Chip Microarray Data

❖ Jonathan Gelfond, The University of North Carolina at Chapel Hill, 823 Hill Top Circle, Sanford, NC 27332, *jgelfond@bios.unc.edu*; Mayetri Gupta, The University of North Carolina at Chapel Hill; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill

Key Words: ChIP-chip, Microarray, Hidden Markov Model, Sequence Analysis

We propose a unified framework for the analysis of Chromatin (Ch)Immunoprecipitation (IP)microarray (ChIP-chip) data and the detection of transcription factor binding sites (TFBSs). ChIP-chip assays are used to focus the genome-wide search for TFBSs by isolating a sample of DNA fragments with TFBSs and applying this sample to a microarray with probes corresponding to tiled segments across the genome. Present analytical methods use the array data to discover peaks or regions of IP enrichment then analyze the sequences of these peaks in a separate procedure to discover the TFBS motifs. The proposed method will jointly model ChIP-chip intensity and DNA sequences through a Bayesian Hidden Markov model (HMM) which identifies TFBSs. The method is applied to simulated and yeast datasets and has favorable TFBS discovery performance compared to current methods.

On Detecting Stabilizing or Divergent Selection Using Patterns of Variation at SNP Loci

❖ Feng Guo, University of Connecticut, 753 C Mix Ave, Hamden, CT 06514, *feng.guo@uconn.edu*; Dipak Dey, University of Connecticut; Kent Holsinger, University of Connecticut

Key Words: Bayesian approach, Hierarchical model, MCMC, SNP, Wright's Fst

We propose several Bayesian hierarchical models to estimate the genetic variations among populations, which is conveniently measured by the Wright's Fst, using single nucleotide polymorphism data from the HapMap project. The posterior distributions of the model parameters are estimated using MCMC simulations and the optimal model is selected using DIC and LPML. To detect loci for which among population variations are not well explained by the common Fst, we use Kullback-Leibler divergence measure (KLD) to measure the divergence between the posterior distributions and the common Fst and calibrate the KLD results using a coin flip experiment. With this method, we identify 15 SNP loci with unusually large values of Fst and we find 10 out of the 15 are located either within identified genes or nearby.

A Bayesian Integrated Approach for Learning About Renal Clear

❖ David Gold, Texas A&M University, 400 Nalge St Apt 406, Statistics Dept., College Station, TX 77840, *dlgold@stat.tamu.edu*; Loleta Harris, The University of Texas M.D. Anderson Cancer Center; Kevin Coombes, The University of Texas M.D. Anderson Cancer Center; Bani Mallick, Texas A&M University

Key Words: microarrays, renal carcinoma, bayesian, hierarchical modeling, gene enrichment

Renal clear cell carcinoma (RCC) is a deadly and complex disease. Attempts to combat RCC would benefit greatly from improvements to the list of candidate genes associated with the disease. Many past microarray studies have failed to identify effective targets for treatment, although more promising results were shown by Lenburg et al. (2003), who compared normal renal to renal tumor gene expression on Affymetrix U133 chips. Identifying effective targets for treatment in high-throughput experiments such as Lenburg et al.'s microarray study is typically complicated by the uncertainty in the gene regulatory networks, i.e. gene interactions, responsible for cancer. We perform multivariate gene inference on the Lenburg et al. microarray data set with a fully Bayesian approach, Bayesian Learning for Microarrays (BLM), with prior information of gene classes.

Using Evolutionary Relationships to Model Correlation in Mixed Effects Models

❖ Hua Guo, University of California, Los Angeles, 3320 Sawtelle Blvd. apt 210, Los Angeles, CA 90066, *guohua@ucla.edu*; Robert E. Weiss, University of California, Los Angeles; Marc A. Suchard, University of California, Los Angeles

Key Words: Bayesian, Evolution, MCMC

Studies of gene expression profiles in response to external perturbation generate repeated measures data that generally follow non-linear curves. To explore the evolution of such profiles across a gene family, we introduce phylogenetic repeated measures (PR) models. These models draw strength from two forms of correlation in the data. Through gene duplication, the family's evolutionary relatedness induces the first form. The second is the correlation across time-points. We borrow a Brownian diffusion process along a given phylogenetic tree to account for the relatedness and co-opt a repeated measures framework to model the latter. We analyze the evolution of gene expression in the yeast kinase family using splines to estimate non-linear behavior across three perturbation experiments. PR models outperform previous approaches and afford the prediction of ancestral expression profiles.

Bayesian Multivariate Spatial Models for Association Mapping in Structured Samples

❖ Meijuan Li, The University of Minnesota, A448 Mayo Bldg., MMC 303, 420 Delaware St. SE, Biostatistics Division, Minneapolis, MN 55455-0378, *meijuanl@biostat.umn.edu*; Brad Carlin, The University of Minnesota; Cavan Reilly, The University of Minnesota

Key Words: Population Structure, Association Mapping, Bayesian Conditional Autoregressive (CAR) Modeling, Linkage-Disequilibrium, Relative Kinship, Deviance Information Criterion (DIC)

Population-based Linkage-Disequilibrium (LD) mapping permits much finer-scale mapping and higher power than does family linkage analysis. However, unlike family based association study, the population-based association mapping is not a controlled experiment, false positives can arise from population structure and genetic familiar relatedness between the samples. There is tremendous interest in simultaneously testing the association between a candidate gene and multiple phenotypes of interest. We here present a new method for population-based association mapping by multivariate Bayesian conditional autoregressive modeling (MCAR). The method we developed accounts for population structure and complex relationships between the samples. We illustrate our modeling approach using the previously published 4 type of flowering data from 95 Arabidopsis thaliana accessions.

60 Statistical Computing and Statistical Graphics Paper Competition ●

Section on Statistical Computing, Section on Statistical Graphics

Sunday, July 29, 4:00 pm–5:50 pm

Improved Centroids Estimation for the Nearest Shrunken Centroid Classifier

❖ Sijian Wang, University of Michigan, 1420 Washington Heights, SPH II, Ann Arbor, MI 48109-2029, *sijwang@umich.edu*; Ji Zhu, University of Michigan

Key Words: High Dimension Low Sample Size, LASSO, Microarray, Nearest Shrunken Centroid, Regularization, Variable Selection

The nearest shrunken centroid (NSC) method has been successfully applied in many DNA microarray classification problems. We show that the NSC method can be interpreted in the framework of LASSO regression. Based on that, we propose two new methods (or penalty functions) for microarray classification, which improve over the NSC. Unlike the L1-norm penalty used in LASSO, the penalty terms that we propose make use of the fact that parameters belonging to one gene should be treated as a natural group. Numerical results indicate that the two new methods tend to remove irrelevant variables more effectively and provide better classification results than the L1-norm approach.

Exploratory Model Analysis with R and GGobi

❖ Hadley Wickham, Iowa State University, 2519 Chamberlain Street, Apt 314, Ames, IA 50014, *h.wickham@gmail.com*

This paper describes exploratory model analysis for ensembles of linear models, where we look at all possible main effects models for a given dataset (or a large subset of these models). This gives greater insight than looking at any small set of best models alone: an ensemble of many models can tell us more about the underlying data than any individual model alone.

spBayes: An R Package for Univariate and Multivariate Hierarchical Point-Referenced Spatial Models

❖ Andrew Finley, The University of Minnesota, Dept. of Forest Resources, 115 Green Hall, 1530 Cleveland Ave. N, Saint Paul, MN 55108, *afinley@stat.umn.edu*; Sudipto Banerjee, The University of Minnesota; Brad Carlin, The University of Minnesota

Key Words: Bayesian inference, coregionalization, kriging, MCMC, multivariate spatial process

Scientists and investigators in diverse fields often encounter spatially referenced data collected over a fixed set of locations within in a region of study. Such point-referenced (geostatistical) data are often best analyzed with Bayesian hierarchical models. Unfortunately, fitting such models involves computationally intensive Markov chain Monte Carlo methods whose efficiency depends upon the specific problem at hand and often requires extensive coding. Here, we introduce a statistical software package, spBayes, built upon the R statistical computing platform that implements a generalized template encompassing a wide variety of Gaussian spatial process models for univariate as well as multivariate point-referenced data. We discuss the algorithms behind our package and software design considerations. We illustrate with an analysis of forest inventory biomass data.

A Flexible Variable Selection Algorithm for the Cox Model with High-Dimensional Data

❖Alexander Pearson, University of Rochester, Department of Biostatistics, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *alexander_pearson@urmc.rochester.edu*; Derick R. Peterson, University of Rochester

Key Words: Model selection, Cox proportional hazards, high-dimensional data

We consider the problem of variable selection with high-dimensional data in the Cox proportional hazards model framework. We propose a search algorithm that lies between forward selection and all subsets. This method uses an evolving subgroup paradigm to intelligently select variables over a larger model space than forward selection. We show that our method can yield significant improvements in the number of true variables selected and the prediction error compared with forward selection, as demonstrated with a simulations study.

61 Open-Ended Questions in Mail, Web, and Telephone Surveys

Section on Survey Research Methods, Section on Government Statistics, SSC, WNAR, Social Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Effects of Answer Space Size on Responses to Open-Ended Questions in Mail Surveys

❖Glenn Israel, University of Florida, PO Box 110540, Gainesville, FL 32611-0540, *gdisrael@ufl.edu*

Key Words: Open-ended questions, mail surveys

Open-ended questions can provide valuable information to help researchers understand a respondent's thinking but using such questions has proven problematic in mail surveys. Experimental data are used to assess the impact of answer space size on the length and content of responses to two open-ended questions that are part of Florida Cooperative Extension's customer satisfaction survey. Based on nearly 2,000 responses collected from 2003 to 2006, the size of the answer space showed no affect on the propensity to answer, but did impact the length of the response with larger spaces eliciting more words and sentences. Results of the content analysis will also be presented. These data provide a context for researchers to make decisions about the use of open-ended questions in mail surveys and the potential impact on the information collected.

Improving the Quality of Open-End Questions in Mail, Web, and Telephone Surveys

❖Don Dillman, Washington State University, 133 Wilson hall SESRC, Pullman, WA 99164-4014, *dillman@wsu.edu*; Jolene Smyth, Washington State University; Leah M. Christian, The University of Georgia

Key Words: open-end questions, survey methods, mail surveys, web surveys, telephone surveys

Results from a series of experiments aimed at improving the quality of open-end questions in mail, web and telephone surveys will be presented. The experiments were conducted annually between 2002 and 2007 using random samples of students (n's approximately 200-350 for each of the experimental panels) at Washington State University, with the aim of testing multiple mechanisms for improving the number of words, number of themes, and extent of elaboration that occur in different modes for the same or similar questions. Results reveal that the quality of open-ended responses can be improved for each of the survey modes, but require somewhat different mechanisms in each mode in order to achieve that goal. The overarching goal for conducting these experiments is to find ways of achieving equivalent quality for open-ended questions across survey modes.

Impact of Presentation Format of Open-Ended Survey Questions on Response Quality

❖Steven Swinford, Montana State University, 2109 Wilson Hall, Bozeman, MT 59717, *swinford@montana.edu*

Key Words: Survey, Open-ended

Six surveys conducted in 2006 on a variety of populations and topics have included variations in the space provided to answer open-ended questions. In each case one-half of the prepared surveys provided white space below the open-ended question to enter a response. In the other half of the surveys the whitespace was framed with a solid line, creating a box, as a visual cue to the respondent as the space where an answer could be entered. Under investigation here are the different response rates using the two formats, the length of the answers provided, whether the borders of the boxed versions were violated when answers were longer than the space provided, and the number of themes contained in answers using both response formats.

The Influence of Color and Space Constraints on Responses to Open-Ended Questions in Paper Questionnaires

❖Virginia Lesser, Oregon State University, Department of Statistics, 44 Kidder Hall, Corvallis, OR 97330, *lesser@science.oregonstate.edu*; Lydia Newton, Oregon State University

Key Words: open-ended questions, mail surveys

A number of experiments were conducted across different populations investigating the use of color and design of the response space allotted for open-ended questions. Eleven studies were conducted from 2004 to 2006. Black and white versions of the questionnaire were compared to either green or blue colored questionnaires. In the visitor surveys, park personnel randomly distributed the questionnaires as individuals entered the park. The size and shape of the area provided for open ended responses varied across the questionnaires in the visitor surveys. For each study, the number of words was determined for each open ended question appearing on the questionnaire. Response rates for the different questionnaire versions were compared. Responses across closed-ended questions were compared to evaluate whether responses differed among the groups.

62 Statistical Methods for Chronic Inflammatory Diseases ● ✪

WNAR, ENAR, Section on Health Policy Statistics

Sunday, July 29, 4:00 pm–5:50 pm

Some Statistical Issues in Rheumatology

❖Charlie Goldsmith, McMaster University, Centre for Evaluation of Medicines, 105 Main Street East Level P1, Hamilton, ON L8N 1G6 Canada, *goldsmit@mcmaster.ca*

Key Words: Design, Factorial, Measurement, Outcome, Validity, Trials

The rheumatic diseases have interesting statistical design issues and innovations in measurement methodology. These include use of experimental designs such as extra-period Latin Squares, Latin Squares, factorial designs and new fractional factorial designs to estimate efficacy when a placebo group is unethical. The Central Limit Theorem has been adapted to make better outcome measurement tools and the recognition of bias in outcome assessment has lead to the development of an "Independent Assessor" as a trained measurer, but not a health professional who is biased in the assessment of these patients; and validating Nurse Measurers to replace Rheumatologists in clinical trials. The innovations have led to improvements in measurement methods and tools that have allowed researchers to become more efficient in the assessment of therapies that can help patients with these diseases.

Meta-Analysis of Clinical Trial Safety Data in Rheumatoid Arthritis (RA) and Other Inflammatory Disease

❖Yun Chon, Amgen Inc., MS 24 2 C, One Amgen Center Dr, Thousand Oaks, CA 91320, *ychon@amgen.com*; Hong A. Xia, Amgen Inc.

Key Words: Rare Adverse Event, Safety Data, Meta-Analysis, Poisson model

Most of clinical trials are designed based on efficacy end points, and therefore have limited power to detect important differences in safety endpoints. The detection of rare AEs is more problematic since studies have relatively short periods and AEs often sparse and leading to zero event in one or both treatment groups. A meta-analysis can be conducted to detect the differences in rare AE between two treatment groups. Results of meta-analysis may be sensitive to the choice of statistical methods. Sensitivity analyses to assess the robustness of results are recommended. The meta-analysis methods at the aggregated study level and at individual patient level using Poisson models and survival techniques are illustrated using clinical trial data. Individual patient data analysis with an appropriate choice of covariates is a preferred method.

Predicting Trends of Radiographic Outcomes Using Baseline Data

❖Grace Park, University of California, Los Angeles, 1025 Yorktown Ave, Division of Biostatistics, Montebello, CA 90640, *gspark@ucla.edu*; Weng-Kee Wong, University of California, Los Angeles; Abdelmonem A. Afifi, University of California, Los Angeles; David A. Elashoff, University of California, Los Angeles; Dinesh Khanna, University of Cincinnati; John T. Sharp, University of Washington; Richard H. Gold, University of California, Los Angeles; Harold E. Paulus, University of California, Los Angeles

Key Words: cluster analysis, K-medians, rheumatoid arthritis, multinomial outcomes

Observational studies may lead to diverse visit times and methods of evaluating radiographic progression of joint damage in rheumatoid arthritis (RA). Instead of traditional progression rates for radiographic outcomes, we examined the usefulness of radiographic profile patterns through clustering algorithms to assess progression rates at set time intervals. Hands and feet radiographic scores were analyzed from a prospective RA cohort. Progression rates were determined by interpolating between set time intervals past the first radiographic observation. Patients were then grouped on their sets of progression rates by clustering algorithms (K-medians) based on Euclidean distances, and cluster membership was examined as a multinomial outcome measure regressed on baseline covariates. We conclude that patterns may correspond better with clinical status and treatment versus traditional rates.

Estimating the Effect of Treatment on Rates of Unintended Events in Rheumatoid Arthritis: A Case History with Caveats

❖Mark Lunt, University of Manchester, Division of PharmacoEpidemiology, 1620 Tremont Street, Boston, MA 02120, *mark.lunt@manchester.ac.uk*; Daniel H. Solomon, Brigham and Women's Hospital; Deborah P.M. Symmons, University of Manchester; Til Stürmer, Brigham and Women's Hospital

Key Words: Propensity Score, effect modification, confounding

Observational studies can provide answers to questions of clinical importance that cannot be answered through a clinical trial. However, such studies can be misleading if imbalances in potential confounders between treated and untreated subjects are not accounted for appropriately. We will present an example of estimating the effect of anti-TNF treatment on mortality in RA patients. A variety of different propensity score (PS) methods were used to balance potential confounders between treated and untreated, giving rise to widely varying estimates of the treatment effect. These different estimates were shown to be due to effect modification across the PS, with different estimators averaging the treatment effect over different distributions of the PS. The results illustrate the importance of assessing effect modification across the PS that can be real or due to unmeasured confounding.

63 Recent Advances in Microarray Data Analysis VI ● ✪

Biometrics Section

Sunday, July 29, 4:00 pm–5:50 pm

Smoothing Spline Mixed Effects Modeling of Multifactorial Gene Expression Profiles

❖Brandon Smith, University of Illinois, 1008 S Busey Ave Apt 6, Urbana, IL 61801, *bjsmith4@uiuc.edu*; Sandra Rodriguez-Zas, University of Illinois at Urbana-Champaign; Bruce Southey, University of Illinois at Urbana-Champaign

Key Words: Smoothing spline, mixed effect model, microarray

The analysis of time-course microarray data is challenging because of the wide range of gene expression patterns, multiple sources of variation, and dependency of the measurements. We studied the performance of a hierarchical approach that integrates a flexible smoothing spline description of time-trends with mixed effects modeling of technical and experimental sources of variation to describe gene expression profiles. Overall patterns of gene expression during honey bee behavioral maturation and deviations associated to race and host colony effects were identified. A penalized likelihood-based approach was used and two spline dimensional bases were compared using likelihood-based criterion. A total of 140 genes exhibited differential expression across age, race and colony. The flexibility of the spline model component permitted the estimation of multiple unique trajectories in time.

Estimating Absolute Transcript Concentration for Microarrays Using Langmuir Adsorption Theory

❖Min Mo, University of Texas at Arlington, Department of Mathematics, Arlington, TX 76019, *minmo@uta.edu*

Key Words: Langmuir adsorption model, Absolute transcript concentration, Microarray, Spike-in genes

The estimation of absolute transcript concentrations might be obtained by finding the estimations of Langmuir parameters. We investigate the possibility of using the spike-in genes found on commercial microarrays, along with an assumed log-linear model for the Langmuir parameters in terms of the spike-in probe sequence features, to estimate the assumed-invariant model coefficients. These estimated coefficients are then used, along with the probe sequence features of the target probes, to estimate the Langmuir parameters for each target probe. Finally, these estimated Langmuir parameters are combined with the expression measurements to produce estimates of the absolute transcript concentrations. The performance of this method will depend on the extent of this extrapolation. Simulation results will be presented to describe the performance of the method.

Positive and Negative Association Rules of Gene Coexpression Patterns

❖Sandra Rodriguez-Zas, University of Illinois at Urbana-Champaign, 1207 W Gregory Dr, Urbana, IL 61801, *rodrgzzs@uiuc.edu*; Bruce Southey, University of Illinois at Urbana-Champaign; Younhee Ko, University of Illinois at Urbana-Champaign; Chengxiang Zhai, University of Illinois at Urbana-Champaign

Key Words: positive and negative association rules, microarray, mixed effects models

The vast amount of data generated from gene expression studies challenges the discovery of meaningful relationships between genes. Interesting associations between genes influencing maturation in honeybees were identified by integrating mixed effects modeling of technical and experimental sources of variation, positive and negative association rules and functional annotation of genes. The highest number of rules occurred when gene-sets were jointly expressed at two or three stages. Multiple interesting rules with high support, high confidence and high conviction were identified. Gene Ontology information identified common molecular function and biological process underlying the co-expression patterns. The synergistic combination of statistical analysis and data mining association rules improved the interpretation of gene expression data.

Model-Based Background Correction for Illumina Microarray Data

❖Yang Xie, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75390-8590, *yang.xie@utsouthwestern.edu*; Lianghao Ding, The University of Texas Southwestern Medical Center at Dallas; Michael Story, The University of Texas Southwestern Medical Center at Dallas; Guanghua Xiao, The University of Texas Southwestern Medical Center at Dallas

Key Words: Illumina Microarray Data, Background Correction, RMA

Illumina Microarray platform become more and more popular due to its high quality and low cost. But the analysis methods for the Illumina array data are far behind than that of the Affymetrix GeneChip data and need to be improved extensively. One feature of illumine arrays is that its noise is controlled by beads conjugated with non-specific oligos. Illumina Inc. provides method to perform background subtraction using the average value of control beads. However, there has been report suggesting that these methods have negative impact on Illumina data quality. Here we propose a new model based method incorporating the negative control beads information into a statistical model for the background correction of Illumina arrays. Both the simulation and real data examples show the good performance of the new method.

An Efficient Mixture Model Approach To Characterize Gene Pathways Using Bayesian Networks

❖Younhee Ko, University of Illinois at Urbana-Champaign, 1207 W Gregory Dr, Urbana, IL 61801, *younko@uiuc.edu*; Sandra Rodriguez-Zas, University of Illinois at Urbana-Champaign; Chengxiang Zhai, University of Illinois at Urbana-Champaign

Key Words: microarray, mixture of models, Bayesian Information Criterion, Bayesian networks

Traditionally, the identification of directional gene networks using Bayesian networks requires the discretization of continuous gene expression measurements that may result in loss of information. A weighted mixture of Gaussian models was used in a Bayesian network to identify gene pathways from continuous gene expression data. Parameter estimates were obtained using an EM algorithm and the network structure better supported by the data was identified using Bayesian Information criterion. To overcome high computational demands, the potential number of genes influencing the expression of another gene was restricted without jeopardizing the parameter search space. The Bayesian network was applied to two published yeast pathways and a honey bee pathway. Results indicated that our approach can efficiently characterize gene pathways with different topology using continuous data.

Penalties Paid for Ignoring the Correlation Between Intensity Channels in Gene-Expression Analysis

❖Sang-Hoon Cho, University of Wisconsin-Madison, MSC 1300 University Ave, Madison, WI 53706, *cho@stat.wisc.edu*; Michael Kosorok, The University of North Carolina at Chapel Hill; Richard A. Johnson, University of Wisconsin-Madison

Key Words: microarray, bivariate gamma distribution, copula function, Bayes estimator

In a typical two channel microarray analysis, it is natural to expect that there exists a strong positive correlation between intensity channels, especially under the null hypothesis of no differentially expressed genes. As far as we are aware, there has not been any study specific investigating about how much gain there would be once the correlation information between intensity channels was utilized in statistical approaches. In this talk, we will show that the correlation information between intensity channels may be a significant component embedded in the data that should not be overlooked in gene-expression analysis. As an illustration, we will show a couple of examples in which we specify a correlation component in statistical methods proposed by Chen et al. (1997) and Newton et al. (2001) and make a comparison between the earlier models and our extended models with correlation.

ANCOVA-Based Normalization Method for cDNA Microarrays

❖Natalia Tchetcherina, The Pennsylvania State University, 326 Thomas Building University Park, State College, PA 16802, *nxt7@psu.edu*; James L. Rosenberger, The Pennsylvania State University

Key Words: normalization, quality control plots, ANCOVA, LSmeans, profile plots, linear model

Normalization is one of important and challenging procedures in the analysis of microarray data. In this paper, we propose using ANCOVA instead of normalization. As a covariate we propose to take the median of spiking control spots or other genes that do not belong to the investigated organism and suppose to be unchanged under treatments. Our method does not have assumption on how many genes are expected to be differentially expressed. The method does not require a fixed number of control spots per array. Even small number of the control spots is enough to use as a covariate. The method can be used for different experiment designs. It can be used for nonlinear models.

64 Longitudinal Data Analysis ● ✪

Biometrics Section, Biopharmaceutical Section, Section on Health Policy Statistics

Sunday, July 29, 4:00 pm–5:50 pm

Inference for Intraclass Correlation Models with Missing Responses at Random

❖ Kai Fun Yu, National Institutes of Health, NICHD/DHH, Bldg 6100, Room 7B05, Bethesda, MD 20892-7510, *yukf@exchange.nih.gov*; Mixia Wu, National Institutes of Health; Aiyi Liu, National Institutes of Health

Key Words: intraclass correlation model, contrast, missing

Intraclass correlation models are popular choices for the treatment of the data from block design or cluster sampling and longitudinal data with an individual random effect. Since missing data often occur in practice, the maximum likelihood (ML) method becomes complicated and requires numerical iterative computations. Some simple exact tests and estimators are often required in some situations. In this paper, we consider an basic intraclass correlation models with missing data at random and propose a new method to construct exact test statistics and simultaneous confidence intervals for all linear contrast in means, which permits missing data with non-monotone pattern. A real example from the Calcium for Preeclampsia Prevention (CPEP) Study is provided for illustration.

Comparing Approaches for Predicting Prostate Cancer from Longitudinal Data

❖ Christopher Morrell, Loyola College in Maryland, Mathematical Sciences Department, 4501 North Charles Street, Baltimore, MD 21210, *chm@loyola.edu*; Larry J. Brant, National Institute on Aging; Shan L. Sheng, National Institute on Aging

Key Words: Classification, Linear Mixed-Effects Models, Sensitivity and Specificity

Classification approaches are compared using longitudinal data to predict the onset of disease. The data are modeled using linear mixed-effects models. Posterior probabilities are computed of group membership starting with the first observation and adding observations until the subject is classified as developing the disease or until the last measurement is used. From the longitudinal analysis we first use the marginal distributions of the mixed-effects models. Next, conditional on group-specific random effects, the conditional distribution is used to compute the posterior probabilities. The third approach uses the distributions of the random effects. Finally, the subjects' data is summarized by the most recent value and rate of change which are used in a logistic regression model to obtain formulae that can be applied at each visit to obtain probabilities of group membership.

Improving the Efficiency of Estimators for Treatment Effect in Two-Arm Randomized Trials Using Auxiliary Covariates

❖ Min Zhang, North Carolina State University, 602 General Joseph Martin Circle, Apt 103, Raleigh, NC 27606, *mzhang4@ncsu.edu*

Key Words: Semiparametrics, Auxiliary Covariates, Randomized Clinical Trials, efficient estimators, longitudinal study

In a randomized clinical trial, the primary goal is to estimate some measure of treatment effect. Because of limited resources, it is desirable to take advantage of whatever information is available to make as efficient inference as possible. In most large trials, many auxiliary covariates are collected, additional information contained in which may form the basis for more efficient inference. Concerns over the potential for bias are often raised over existing methods for covariate-adjusted analysis. Moreover, such "adjusted" analysis estimate treatment effect conditional on covariates, whereas the primary interest often focuses on the unconditional effect. Using the theory of semiparametrics, we propose a general and flexible approach to exploit auxiliary covariates that leads to more precise estimators for unconditional treatment effect while circumventing the usual concerns.

Moving Block Bootstrap for Analyzing Longitudinal Data

❖ Hyunsu Ju, The University of Texas Medical Branch, 3.302 Jennie Sealy Hospital 301 University Blvd, Galveston, TX 77550, *hyju@utmb.edu*; Suojin Wang, Texas A&M University

Key Words: Longitudinal study, Resampling methods, Moving block bootstrap

Many longitudinal studies in which subjects are followed over time are the case that the number of subjects is large and the number of measurements per subject over time is small. We focus on a case where the number of replications over time is large relative to the number of subjects in the study. We investigate the use of moving block bootstrap methods involving resampling possibly overlapping blocks for analyzing such data rather than standard ordinary bootstrap or resampling subjects bootstrap. Asymptotic properties of the bootstrap methods in this setting are derived. The effectiveness of these block resampling methods is also demonstrated through a simulation study.

Comparison of Design-Based and Model-Based Methods to Estimate the Cluster Effect Using National Population Health Survey Data (1994–2003)

❖ Sunita Ghosh, University of Saskatchewan, 1042711 112th St, Edmonton, AB T6J 4M1 Canada, *sug132@mail.usask.ca*; Punam Pahwa, University of Saskatchewan

Key Words: Multi-stage sampling, Bootstrap, Generalized Estimating Equation, NPHS, Asthma, Logistic Regression

Analysis of data collected using multi-stage sampling design, should account for stratification, clustering and unequal probability of selection. This paper aims to account for clustering and unequal probability of selection. Design-based approaches Rao-Wu bootstrap methods were used. Generalized Estimating Equation (GEE) approach was used for model-based methods. Both these approaches were applied and compared using National Population Health Survey (NPHS) dataset. Two time points Cycle 1 (1994–95) and Cycle 5 (2002–03) were used for analytical purpose. Longitudinal weight provided by Statistics Canada was used. The variable of interest was self-reported physician diagnosed asthma. Its relationship with other variables was studied. Accounting for the two features of survey design, produced similar results for the design-based and model-based methods.

Highly Efficient Longitudinal Designs That Are Robust Against Mis-specification of the Model

❖ Shirley Ortega Azurduy, Maastricht University, PO Box 616, Maastricht, NL-6200 MD Netherlands, *shirley.ortega@stat.unimaas.nl*; Frans Tan, Maastricht University; Martijn P.F. Berger, University of Maastricht

Key Words: D-optimality, linear mixed models, longitudinal data, Maximin criterion, robust design

Longitudinal data are frequently analyzed by linear mixed models. Optimal designs for these analyses, however, depend on the specified model. Practitioners seldom know the model structure before data collection. When the underlying true model differs from the model that is used to specify the optimal design, this design becomes inefficient. We apply a D-optimal maximin criterion to find a set of designs that remain highly efficient whenever the model is incorrectly specified. Based on different model and parameter spaces, the relative efficiency of each chosen optimal design with respect to the true model is computed. Thereafter, the minimum relative efficiencies as function of the true models are derived. Finally, designs with the maximum efficiencies are identified. We show that these robust designs have a larger efficiency than the conventional equally spaced designs.

Solving Structured Maximum Likelihood Estimating Equations for Nested Marginal Discrete Longitudinal Models

❖ Arunava Chakravartty, University of California, Riverside, 200 W Big Springs Road, Apt 4, Riverside, CA 92507, *achak001@ucr.edu*

Key Words: estimating equations, marginal models, longitudinal data, nested models, numerical convergence, intial solution

Three nested marginal models are considered for describing a discrete longitudinal data. The Maximum Likelihood Estimating Equations for these three models have a common set of Estimating Equations with a special structure. This paper brings forth the structure common to these set of Estimating Equations and then describes the methods for solving the estimating equations for all three nested models. The initial solutions in these methods are determined naturally from the Estimating Equations and then improved to the final solutions through a series of steps described in the paper.

65 Statistical Issues Related to Categorical Data in Clinical Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Sunday, July 29, 4:00 pm–5:50 pm

Sample Size Determination for Binomial Endpoints Based on the Width of the Confidence Interval in Phase IIA Clinical Trials

❖ Tai-Tsang Chen, Bristol-Myers Squibb Company, 5 Research Parkway, 2FW311, Wallingford, CT 06492, *tai-tsang.chen@bms.com*; Vishwanath Iyer, Novartis Healthcare Pvt. Ltd; Aparna B. Anderson, Bristol-Myers Squibb Company

Key Words: Sample size, Confidence interval, phase II

Single-arm phase II clinical trials are usually undertaken to assess whether a new agent has sufficient activity and an acceptable safety profile to warrant further investigation. When the goal is to estimate the response rate, the design options are limited. We propose an approach to determine the sample size in a phase II study when the primary interest is to estimate a binary outcome. The sample size is chosen to ensure that among confidence intervals that contain the true parameter, the probability that the interval width is smaller than a prespecified threshold is sufficiently high. We will apply the sample size calculation to different methods of calculating confidence intervals and provide recommendations on how to choose the threshold value for the confidence interval width.

Sample Size Estimation for the Phase III Confirmatory Clinical Trials Considering the Uncertainty of the Effect Size

❖ Zhe Shang, Novartis Pharmaceuticals, One Health Plaza, East Hanover, 07936, *zhe.shang@novartis.com*; Yawei Zhang, University of Pennsylvania

Key Words: confirmatory clinical trials, sample size, effect size

Sample size estimation is a crucial component during the designing stage of phase III confirmatory clinical trials. Wang, Hung, and O'Neill (2004) pointed out that the estimates of effect sizes obtained from previous clinical trials and past literatures are often too optimistic for the planning of the phase III confirmatory trials. Sample size estimation considering the uncertainty of the variance was studied by Julious and Owen (2006). In the presentation, we developed three sample size estimation formulas which take the uncertainty of effect size into consideration. We discussed the properties of our sample size estimation formulas and compared them with the traditional sample size calculation formula. We also pointed out how our sample size formulas can be applied to the binary response data. Finally we applied our sample size estimation formulas to a real phase III clinical trial.

Small-Sample Tests for Efficacy in Fixed-Dose Drug Combination Studies

❖ Melinda M. Holt, Sam Houston State University, Mathematics and Statistics, PO Box 2206, Huntsville, TX 77341, *mholt@shsu.edu*; Jesse Linder, Southeastern Louisiana University; Ferry Butar Butar, Sam Houston State University

Key Words: Binomial, Clinical Trials, Resampling, Bootstrap, Min Test

To be approved by the U.S. Food and Drug Administration, a drug combination must demonstrate efficacy; that is, it must be superior to each of its components. Here we consider fixed-dose clinical trials designed to show the efficacy of a combination drug when the response of interest is binary. Wang and Hung (1997) developed large-sample tests to demonstrate efficacy. Based on the results of a performance analysis, the authors recommended their tests for samples of size 20 or more and called for further study into small-sample methods. We present small-sample procedures based on the bootstrap and permutation methods. We also provide a performance analysis comparing the power and Type I error of the small-sample approaches to those of the large-sample approach.

Extension of the Rank Sum Test for Clustered Data in Vision Clinical Trials

❖ Ratna Ramaswamy, Johnson & Johnson PRD, 199 Grandview Road, Skillman, NJ 08502, *rramaswa@cpcus.jnj.com*; Gary Koch, The University of North Carolina at Chapel Hill; Ram Ganeshan, Statistical Consultant

Key Words: Wilcoxon Type Nonparametric Rank Tests, Vision Trials, Non-Independence among experimental Units

In Vision Clinical trials, the unit of drug treatment is one of the two eyes. To make the maximum use of the experimental units eye is considered as unit of analysis. Only the case of two treatments, active and vehicle placebo is considered here for simplicity. Clinical parameters of interest as an example would be itching and redness measured on the ordinal scale. For the nature of the data (generally, non-normal and skew distributed), a rank test such as two-sample Wilcoxon Rank Sum Test comes out as the test of choice. However the factor of non-independence of the sampling units (eyes) is generally ignored. Some reasonable ranking methods using Cochran-Mantel-Haenzel tests that would take the correlated structure of the data into consideration are given here.

Estimation of Diagnostic Accuracy Measures for a Binocular Test

❖Daniel Bonzo, Altus Pharmaceuticals, Inc., 125 Sidney Street, Cambridge, MA 02703, *dbonzo@altus.com*; Alexander R. de Leon, University of Calgary

Key Words: Common correlation model, correlated binary data, coverage probability, predictive values, sensitivity, specificity

Diagnostic studies in ophthalmology frequently involve binocular data where pairs of eyes are evaluated, through some diagnostic procedure, for the presence of certain diseases or pathologies. The simplest approach of estimating measures of diagnostic accuracy, such as sensitivity and specificity, treats eyes as independent. Approaches which account for the inter-eye correlation include regression methods using GEE and likelihood techniques based on various correlated binomial models. The paper proposes a simple alternative methodology of estimating measures of diagnostic accuracy for binocular tests based on a flexible model for correlated binary data. Moments estimation of model parameters is outlined and asymptotic inference is discussed. The computation of the estimators and their standard errors are illustrated with data from a study on diabetic retinopathy.

Exposure Adjusted Incidence and Event Rates in Safety Analysis

❖Li Chen, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, *lic@amgen.com*; Xin He, University of Missouri-Columbia; Lei Lei, Amgen Inc.; Hong A. Xia, Amgen Inc.

Key Words: Adverse event, Crude percentage, Exposure adjusted incidence rate, Exposure adjusted event rate, Confidence interval

Assessment of drug safety typically involves estimation of occurrence rate of adverse events. Most often, the crude percentage (subject incidence) is used to estimate adverse event rate. However, in some situations, the exposure adjusted incidence rate (EAIR) and exposure adjusted event rate (EAER) may be more appropriate measures to account for the potential difference in the duration of drug exposure or the follow-up time among individuals, and to capture the multiple occurrences of certain adverse events for a subject. In this paper, we establish the asymptotic properties of the EAIR and EAER under certain assumptions, and propose a general approach for variance estimation of the rates and for calculating the confidence intervals of the rate differences between two groups. The methodology is illustrated by simulation studies and applications in a clinical trail.

Conditional Confidence Intervals for p1-p2 and p1/p2 Using a Generalized Fisher's Exact Test

Devan V. Mehrotra, Merck & Co., Inc.; ❖Arthur Roth, Pfizer Inc., 2800 Plymouth Rd, Ann Arbor, MI 48105, *arthur.j.roth@pfizer.com*

Key Words: nuisance parameter, confidence interval, Fisher's exact test, difference of proportions, ratio of proportions

Fisher's exact test compares two independent binomial proportions, p1 and p2, using conditional inference. It can be viewed as a test on the odds ratio OR=[p1/(1-p1)]/[p2/(1-p2)], the difference D=p1-p2, or the ratio R=p1/p2, i.e., the null hypothesis can be written as OR=1, D=0, or R=1. A corresponding exact conditional confidence interval (CI) to complement the p-value from Fisher's test is easy to construct for OR based on a natural ancillary statistic, but not for D or R, which have no ancillary statistics, making it necessary to deal with a nuisance parameter. We propose two solutions to the nuisance parameter dilemma, each within the context of a "generalized Fisher's exact" (GFE) test. Illustrative examples and exact numerical results are provided to demonstrate the superiority of our GFE-based conditional CIs for D and R relative those proposed by Santner and Snell (JASA, 1980).

66 Use of Bayesian Statistics in Phase-I Clinical Studies ● ✪

Biopharmaceutical Section, Biometrics Section

Sunday, July 29, 4:00 pm–5:50 pm

Maximum Likelihood Methods for Optimal Treatment Schedule-Finding in Clinical Trials

❖Changying Liu, University of Michigan, 38175 Lantern Hill court, Farmington Hills, MI 48331, *changying8618@yahoo.com*; Thomas M. Braun, University of Michigan

Key Words: Phase I trial, Maximum tolerate schedule, change-point, boundary parameter, maximum likelihood, large sample property

Braun, Yuan, and Thall (2005) constructed a Bayesian paradigm for Phase I trial designs that allows for the evaluation and comparison of several treatment schedules, each consisting of a sequence of administration times. The goal of their design was to determine a maximum tolerated schedule (MTS). Our work focuses upon a maximum-likelihood alternative to their work. We propose using maximum likelihood to estimate the parameters of the triangular hazard used by Braun, Yuan, and Thall (2005) in a single administration setting. We describe how to derive estimators for the change-point and boundary parameters of the triangular hazard model and discuss their large sample properties. We also explain the challenge in extending the results from a single administration setting to a multiple schedules setting, with each schedule consisting of multiple administrations.

Power and Sample Size for Dose Proportionality Studies

❖Ilsung Chang, Johnson & Johnson PRD, 900 Davidson Road, APT 145, Piscataway, NJ 08854, *ichang@prdus.jnj.com*; Jaya Natarajan, Johnson & Johnson PRD

Key Words: Dose proportionality, Power and sample size, Pharmacokinetics

The pharmacokinetics of a compound are defined to be dose-proportional if the AUC and Cmax are proportional to the dose. The analysis of dose proportionality studies can be handled in several ways: ANOVA, confidence intervals for each pair of doses studied (BE approach), or confidence interval on the estimated slope of the regression line. In this presentation, for crossover and parallel group designs, we examine the power and sample size required for the two confidence interval approaches (CI for each pair of doses; CI for the estimated slope) using both closed form expression and simulation approach.

Elicitation of Expert Opinion and Bayesian Fitting of a Logistic Dose-Response Curve

❖Les Huson, GNB Limited, Chislehurst Business Centre, Bromley Lane, Chislehurst, BR7 6LH United Kingdom, *les@gnblimited.co.uk*; Nelson Kinnersley, Roche Products Limited

Key Words: bayesian, logistic, dose-response, elicitation, prior

Before undertaking a Phase I study in which a dose-response relationship was to be investigated, we elicited prior opinion about the relationship from a group of internal experts, using a graphical technique. We describe this elicitation process, and also discuss two methods which we used to convert the expert opinion into prior distributions for the parameters of a logistic dose-response curve. The first such technique combines the CDF approach with a simulation step, and the second involves forming priors by bootstrapping. We also report the results from Bayesian analyses combining these priors with data from the Phase I study.

A Discussion on Operational Characteristics in Bayesian Clinical Trials

❖ Xin Zhao, Merck & Co., Inc., 150 Oakland Ave Apt A307, Lansdale, PA 19446, *xin_zhao@merck.com*

Key Words: Bayesian clinical trial, operational characteristics, Bayesian methods, clinical trial

Bayesian methods have been applied in clinical trials in recent years. Operational characteristics are very important in Bayesian clinical trial planning and evaluation. In this talk, I will lay out the aspects we need to consider when we plan the operational characteristics in a Bayesian clinical trial. I will also discuss how to evaluate a Bayesian clinical trial results based on its operational characteristics.

EWOC Online: An Interactive Web Tool for Dose Escalation in Cancer Phase I Clinical Trials

❖ Zhiheng Xu, Emory University, 1365B Clifton Rd, Room B4112, Atlanta, GA 30322, *zhiheng.xu@emoryhealthcare.org*; Haibin Wang, Emory University; Mourad Tighiouart, Emory University; André Rogatko, Emory University

Key Words: EWOC, Maximum tolerated dose, Bayesian methods, Cancer phase I clinical trial

EWOC (escalation with overdose control) online is a web-based program which uses Bayesian methods in dose escalation schemes to assign dose levels sequentially to patients in cancer phase I clinical trials. It has a detailed tutorial which guides users in designing phase I trial with the goal of approaching maximum tolerated dose as fast as possible while controlling the probability of overdosing patients. Table for sequential dosages and plot for tree of doses can be generated from the program. The dose assignment can be recalculated by modifying the input parameters. Simulation module will also be available for users to simulate a real phase I trial for a certain number of cancer patients. Operating characteristics such as consistency and safety of the trial using this methodology have been well established. EWOC online can be accessed at http://sisyphus.emory.edu:8080/ewoc/ewoc3.htm.

Semiparametric Mixed Analysis on PK/PD Models Using Differential Equations

❖ Yi Wang, University of Nebraska-Lincoln, 4300 Holdrege Street A201, Lincoln, NE 68503, *wangyi9876@yahoo.com*; Kent Eskridge, University of Nebraska-Lincoln; Shunpu Zhang, University of Nebraska-Lincoln

Key Words: semiparametric, ODEs, spline-enhanced, nonlinear mixed models

Motivated by the use of semiparametric nonlinear mixed-effects modeling on longitudinal data, we develop a new ordinary differential equations (ODEs) based method designed for general PK/PD models, which are not addressed by Ke and Wang (2001) and Li et al. (2002). In our method the part of ODEs is relaxed and assumed to be nonparametric functions that will be decided by data and estimated nonparametrically using splines. This setup makes identification of structural model misspecification feasible by quantifying the model uncertainty and provides flexibility for accommodating possible structural model deficiencies, which are accomplished by implementing the spline-enhanced differential equations in a nonlinear mixed-effects modeling setup. We illustrate the method with an application to cefamandole data and we evaluate its performance through simulation.

67 Longitudinal Data Analysis ●

ENAR, Section on Health Policy Statistics, Biometrics Section
Sunday, July 29, 4:00 pm–5:50 pm

Pool-Testing Samples with Random Intercept

❖ Peng Chen, University of South Carolina, 100 Riverbend Drive, Apartment G24, West Columbia, SC 29169, *chenp@stat.sc.edu*; Joshua M. Tebbs, University of South Carolina

Key Words: pool testing, random intercept, MCEM algorithm, GLMM

In this paper, we extended the regression model discussed in Vansteelant etc. (2000) by introducing a random site effect. The maximum likelihood estimate of parameters are discussed. Both direct maximization via numerical qudarartures and Monte Carlo EM algorithm are proposed. Variance-covariance matrix for MLE are computed, a likelihood-based test for random effect are derived.

Multiple Parameter Combination of Between-Cluster and Within-Cluster Regression Coefficient Estimates

❖ Julie A. Stoner, University of Nebraska Medical Center, Division of Biostatistics CoPH, 984350 Nebraska Medical Center, Omaha, NE 68198-4350, *jstoner@unmc.edu*; Brian G. Leroux, University of Washington

Key Words: Between-cluster Effects, Within-cluster Effects, Correlated Data, Generalized Estimating Equations

A marginal regression modeling method that optimally weights and combines between-cluster and within-cluster coefficient estimates in a correlated data setting is presented. The combination accounts for correlation among multiple between-cluster and within-cluster coefficient estimators and may be applied to multiple coefficient parameters in a given model. In settings where combining coefficient estimates is appropriate, the method results in increased efficiency while maintaining stability and coverage probabilities and avoids direct correlation parameter estimation. The method is appealing, relative to typical implementations of GEE, in settings with large cluster sizes relative to the number of independent clusters, varying cluster sizes, and complicated correlation structures. Method performance is studied by simulation and an application. (NIH grant R01DE15651)

Combined Linkage and Association Mapping of Quantitative Trait Loci with Missing Genotype Data

Ruzong Fan, Texas A&M University; ❖ Lian Liu, Texas A&M University, Dept of Statistics, College Station, TX 77843, *lian@stat.tamu.edu*; Jeesun Jung, Indiana University; Ming Zhong, Texas A&M University

Key Words: Missing genotype, Linkage disequilibrium mapping, QTL

In this paper, the impact of missing genotypes is investigated for high-resolution combined linkage and association mapping of quantitative trait loci (QTL). We assume that the genotype data are missing completely at random (MCAR). Based on two regression models, F-test statistics are proposed to test association between the QTL and markers. The non-centrality parameter approximations of F-test statistics are derived to make power calculation and comparison, which show that the power of the F-tests is reduced due to the missingness. By simulation study, we show that the two models have reasonable type I error rates for a dataset of moderate sample size. As a practical example, the method is applied to analyze the angiotensin-1 converting enzyme (ACE) data.

Analyzing Correlated Data with Partial and Small Clusters

❖Xianqun Luan, The Children's Hospital of Philadelphia, 34th St and Civic Ctr Blvd, CHOP North, Suite 1457, Philadelphia, PA 19104, *luan@email.chop.edu*; Avital Cnaan, The Children's Hospital of Philadelphia

Key Words: Multiple outputation, Generalized estimating equations, within-cluster resampling

Studies in clinical research may have a setup where only a small proportion of the study sample has clusters, and those clusters are small. For example, studies in children when all siblings within an age range are in the study. One may use correlated analysis techniques: mixed effects models or generalized estimating equations. However, these approaches generally assume that most of the data are clustered, and that the clusters mostly are of size n>2. An alternative approach is multiple outputation or within-cluster resampling. We compare these various methods with simulation studies for binary outcomes. We vary: percentage of clustered data; size and patterns of the actual clusters; correlation within the cluster; and proportion of successes under H0 and H1. We explore cluster size correlated with outcome. We define guidelines of when to use which approach based on these simulations.

Using the Power of an F-Test To Compare Designs for Response Surface Models with Random Block Effects

❖Sourish Saha, University of Florida, 1001 SW 16th Avenue, Dept of Statistics, Apt 117, Gainesville, FL 32601, *sourish@ufl.edu*; Andre I. Khuri, University of Florida

Key Words: linear mixed models, prediction variance, quantile dispersion graphs, power of an F test, response surface designs

One traditional criterion for comparing response surface designs is the use of the prediction variance. Another criterion is based on utilizing the power of an F-test concerning the fixed unknown parameters in the associated model. The purpose of the talk is to compare response surface designs on the basis of the power criterion when the response surface model contains a random block effect, in addition to the fixed polynomial portion of the model. The proposed approach uses quantiles of either the power function or the corresponding noncentrality parameter on concentric surfaces within the so-called alternative space associated with the unknown parameters. The dependence of these quantiles on the unknown value of the ratio of two variance components, namely, the ones for the block effect and the experimental error, is depicted by plotting the so-called quantile dispersion (QDGs).

Minimum Hellinger Distance Estimation of Finite Mixtures of Generalized Linear Models

❖Jing Shen, IBM T.J. Watson Research Center, 71 Charter circle, Apt 4A, Ossining, NY 10562, *jingshen@us.ibm.com*; Daniel Hall, University of Georgia; Chenhua Zhang, University of Georgia

Key Words: Conditional Density, Robustness, Mixture Models, Zero Inflated Poisson, Zero Inflated Binomial

We study minimum Hellinger distance estimation (MHDE) method for finite mixtures of generalized linear models (FMGLMs). MHDE method as a robust approach is well established for iid case. Recently, it has been extended to mixture models with regression context but leaving asymptotic properties unexamined. In this paper, we define a new MHDE method in the general FMGLMs context based on conditional densities. We prove that our method yields consistent and asymptotic normally estimators. Numerical results suggest that our method is more robust than MLE with the presence of outliers and competitive otherwise. In addition, we demonstrate that our method is more efficient than the previous one in the literature and can be applied more broadly. An example on a cohort study which characterizes ambulatory electrocardiography results of overtly healthy dog is used to illustrate our method.

Design and Analysis of Multiple Control Arm Studies

❖Laura Johnson, National Institutes of Health, NCCAM, 6707 Democracy Blvd, Suite 401 MSC 5475, Bethesda, MD 20892-5475, *johnslau@mail.nih.gov*; Catherine M. Stoney, National Institutes of Health

Key Words: mixed models, control groups, placebo, research design

Placebo controlled studies are often considered impossible or unethical for certain populations and topics under study. Thus, alternative treatment or other control groups are often included in designs to better characterize the effects of experimental treatments. However, it is clear from specific investigation of placebo responding that the placebo response plays a measurable role in clinical research. How and which control groups are included in experimental designs greatly influences choice of analysis and what information can be ultimately derived from the analyses. We will focus on different types of control groups, limitations and advantages of placebo control groups, why more than one control group may be needed in a study even when not dismantling a complex intervention, and how choice of analysis model can give different insights into the same study.

68 Hardware, Software, and Algorithms ● ✪

Section on Statistical Computing, Section on Statistical Graphics

Sunday, July 29, 4:00 pm–5:50 pm

WISDOM for μStat: Web-Based Support for the Analysis of Multivariate Hierarchical Data

❖Knut M. Wittkowski, The Rockefeller University, Cntr for Clinical and Translational Sc, 1230 York Ave Box 322, New York, NY 10021, *kmw@rockefeller.edu*

Key Words: non-parametric, u-statistics, multivariate, factorial, knowledge-based, interactive

As the amount of genetic, genomic, proteomic, and phenomic data is increasing, complexity of data management and analysis are also increasing. At the CCTS, we are approaching the new challenges by (1) extending u-scores to multivariate data, including specific partial orderings (e.g., for genomic data), (2) building a grid of PCs to handle whole genome scans, and (3) generating centralized, secure, and Web accessible data bases from a interactive graphical description of the research design. We are now extending this Web-based system (WISDOM) to supporting our clinical investigators with statistical analyses. In assisted mode (confirmatory analyses), statistical analyses are essentially automated, based upon the knowledge about the primary aims. In expert mode (exploratory analyses), the same knowledge base is used to provide study-specific defaults.

Creating Statistical Web Services Using ASP.NET

❖Neil Polhemus, StatPoint, Inc., 11133 Victor Drive, Marshall, VA 20115, *neil@statpoint.com*

Key Words: software, graphics, XML, .NET, web services, scripts

ASP.NET XML Web Services provide a useful mechanism for exposing statistical calculations and graphics to web application developers. Using this approach, the capabilities of statistical software packages can be tapped by developers without invoking the software's GUI. Data and instructions are passed to the package using XML files. Output (including graphics) is returned as HTML with imbedded images, while numerical results are made available to the calling program in XML files. The same scripts that control the web services can also be accessed by the software package's user interface to reproduce the same analyses, or the GUI can be used to help create the XML scripts. This talk examines some of the issues involved in implementing such an approach, using the STATGRAPHICS statistical software package as an example.

Grid Computing

❖Abdullah Alnoshan, George Washington University, 344 Maple Avenue West, 321, Vienna, VA 22180, *noshan@gwu.edu*; Shmuel Rotenstreich, George Washington University; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University; Adil Rajput, BearingPoint

Key Words: distributed systems, Resource allocation, peer-to-peer grid, utility function

Research advances in the distributed computing has led to Grid computing, which is unlike many other areas of distributed computing focuses on large-scale resource sharing and discovery for better performance. Traditionally, Grid computing has been performed in a client/server format. However, recent advances in the area of peer-to-peer computing have brought a paradigm shift in the field of distributed systems and have forced researchers in the area of Grid computing to discuss peer-to-peer grids. We will review the work done in the area of peer-to-peer Grids and will provide the first step in defining a framework that provides incentive to the participating resources in a peer-to-peer grid environment. The framework provides the basis for defining a utility function that is used by participating nodes to efficiently allocate the resources that are available in a system.

Access Control Model for E-Learning System

❖Fahad Bin Muhaya, Imam University, PO Box 84901, Riyadh, 11681 Saudi Arabia, *gmuva@hotmail.com*; Yasmin H. Said, George Mason University

Key Words: Information System, Information Technology, Educational System, e-learning management system, Access Control

The educational system has gradually shifted from a face-to-face to an e-learning system, which has become prevalent in advanced countries with the advance of information technology, and connection of global networks. Therefore, in this paper, we develop an access control model particularly for e-learning management system.

A Web-Based Program for Computing Percentage Points of Pearson Distributions

❖Wei Pan, University of Cincinnati, PO Box 210049, Cincinnati, OH 45221, *wei.pan@uc.edu*; Haiyan Bai, University of Central Florida; Shengbao Chen, JMW Truss & Components

Key Words: curve fitting, Pearson distributions, percentage points, web-based program

The applications of curve fitting using Pearson distributions have called for an efficient program to compute percentage points of Pearson distributions. Endeavors have been made by previous researchers; however, all of the existing computing programs are locally dependent on a certain statistical software. The current study develops a web-based program to facilitate researchers to compute the percentage points. The end-users do not need to install statistical software on their local computers, which saves the hassle of debugging the coding. The user-friendly interface allows the users to input parameters and obtain the estimates by simply hitting a "Run" button. The validity of the parameters is automatically checked before computing. The results of the computation with a dynamic probability curve are presented directly on the screen for printing, coping, and saving.

minSpline: An R Package for Fitting Splines

❖Sundar Dorai-Raj, PDF Solutions, Inc., 333 W San Carlos, Suite 700, San Jose, CA 95110, *sundar.dorai-raj@pdf.com*; Spencer Graves, PDF Solutions, Inc.

Key Words: smoothing, curve, surface, fitting

DIERCKX is a package of Fortran subroutines for calculating smoothing splines for various kinds of data and geometries, with automatic knot selection. We have created a package for accessing these functions in R. Our presentation will cover current spline implementations in R and how the new "minSpline" package extends beyond the current work.

Calculating the Interatomic Distance Distribution from Small-Angle X-Ray Scattering via Curve Averaging

❖Lanqing Hua, Purdue University, 2108 W. White Street, Apt. 141, Champaign, IL 61821, *lhua@stat.purdue.edu*; Alan Friedman, Purdue University; Chris Bailey-Kellogg, Dartmouth University; Bruce Craig, Purdue University

Key Words: indirect transform, model averaging

Small-angle X-ray scattering (SAXS) is a useful technology to study a molecule's structure. Relative to other technologies, it provides structural information with a reduced investment in time and effort. Reconstruction of the interatomic distance distribution, P(r) curve, has been a long standing problem with SAXS experiments. In this paper, we present an indirect transform method using "model averaging". This approach expands on Moore's (1980) method by using multiple sets of basis functions to fit the observed scattering data and reconstructing the P(r) curve as a weighted average of these different representations. This approach completely avoids the need to impose subjective physical assumptions about the protein or restraints on P(r), common with other approaches. Through simulation, we show that our method provides comparable or more accurate P(r) estimation than Moore's method.

69 Section on Statistics in Sports Contributed ●

Section on Statistics in Sports

Sunday, July 29, 4:00 pm–5:50 pm

The Distribution and Simulation of Baseball Winning Percentages

❖Richard Auer, Loyola College in Maryland, Loyola College in Maryland, 4501 North Charles Street, Baltimore, MD 21210, *rea@loyola.edu*

Key Words: Major League Baseball, Multivariate Normal Distribution, Simulation, Conditional Normal Densities

Many major league baseball purists state that the wild card team receives an easy ride to the playoffs. After all, the wild card team is not even the champion of their own division. Using real major league baseball data and modeling a single season's worth of winning percentages as a multivariate normal random vector (with first place to last as the variates), a simulation study suggests that that the wild card team is typically much stronger than the weakest of the divisional winners. A goodness of fit study of the multivariate model is described along with a detailed look at how conditional normal densities are used to simulate many seasons of data, one place of finish at a time.

Assessing Pitcher and Catcher Influences on Base Stealing in Major League Baseball

❖Thomas M. Loughin, Simon Fraser University, Department of Statistics and Actuarial Science, 250 - 13450 102nd Avenue, Surrey, BC V3T0A3 Canada, *tloughin@sfu.ca*; Jason Bargen, Sprint Nextel

Key Words: mixed model, logistic, sports, pitchers, catchers

A formal statistical analysis is performed on play-by-play data to determine the extent to which pitchers and catchers can influence stolen-base attempts and successes. Two response proportions—attempt/opportunity and success/attempt—are modeled separately using mixed-effects logistic regression. Pitchers and catchers are entered as random effects and various other factors thought to influence stolen-base attempts and successes are entered as fixed effects. Variance components are estimated and hypotheses tests indicate that the population variance components for both pitchers and catchers are significant for both response proportions. The presence of variation among players at the respective positions is interpreted as evidence that stolen-base defense is a real skill exhibited to varying degrees by different players.

Sorting Algorithms and Sport Tournaments

❖Hans Block, KTH, Stockholm, Sweden, Monstringsvagen 126, Akersberga, S-18433 Sweden, *hans.block@telia.com*

Key Words: Parallel sorting, Sport tournaments

Many types of tournament schemes are used in different sports. We want to construct a scheme which gives the whole ranking order when pairs of athletes or teams compete. The scheme should be efficient, fair, and thrilling. The problem can be formulated as parallel sorting. The proposed algorithm will, according to simulations, take approximately $2.4\cdot\lg n$ rounds in a tournament with n participants. The computations will in most cases take $O(n^{2}\cdot\log\ n)$ steps. However, both faster and slower special cases exist. Some limits for the number of rounds are given for variants of the algorithm. The scheme has been used with good results in real tournaments with up to 40 participants. At last we discuss in which branches of sports this algorithm will be most useful.

Batter-Pitcher Matchups in Baseball

❖Hal Stern, University of California, Irvine, Irvine, CA 92612, *sternh@uci.edu*; Adam Sugano, University of California, Los Angeles

Key Words: hierarchical model, beta-binomial distribution, baseball statistics

Baseball managers often make decisions about which players to use in a game based on the results of previous matchups with a given pitcher. These previous results are typically based on small samples. A hierarchical beta-binomial model is used to assess the variability in batter performance across the population of pitchers and to assess the variability of pitcher performance across the population of batters.

Statistical Analysis in One-Day Cricket

❖Wathurawa A. Bandulasiri, Sam Houston State University, Dept of Mathematics Statistics, Huntsville, TX 77340, *abandulasiri@shsu.edu*; Ferry Butar Butar, Sam Houston State University

Key Words: Cricket, Logistic Regression, Sports, Odds Ratio

Winning a one day international (ODI) cricket match could depend on various factors related to the strengths of the two teams. While some of these factors have been analyzed and well documented in the literature, some are yet to be investigated. Based on the analysis in this paper, statistical significance of a range of variables that could explain the outcome of an ODI cricket match are explored. In particular, home ground advantage, winning the toss, game plan (batting first or fielding first with the winning of the toss), match type (day or day and night) and the effect of the Duckworth-Lewis method for the matches shortened due to rain will be key interests in the investigation.

A Bayesian Approach to Ranking Women's College Hockey Teams

❖Michael Rutter, The Pennsylvania State University, Erie, The Behrend College, 4205 College Dr, Erie, PA 16563-0203, *mar36@psu.edu*

Key Words: Sports, Ranking, Paired Comparisons, Bayesian

A key criterion used to determine the participants of the NCAA Women's Ice Hockey Championship tournament is a team's Rating Percentage Index. This rating system only takes into account a team's overall and its opponent's record, not game results against a specific opponent. Based on Mease's work, I propose a ranking system based on the results of individual games, ignoring game score. In women's hockey, approximately 10% of regular season games end in ties, and the proposed ranking system explicitly models the probability of two teams tying, given the relative strength of each team. A Bayesian approach is utilized, allowing for information on the probability of a tie from previous seasons to be included in the model. In addition, the prior distribution of a team's rating can be based on previous year's results, allowing for early season rankings to be indicative of the final rankings.

Bivariate Score Plots

❖Rafe Donahue, Vanderbilt University, Medical Center, Nashville, TN 37232, *rafe.donahue@vanderbilt.edu*; Tatsuki Koyama, Vanderbilt University; Jeff Horner, Vanderbilt University; Cole Beck, Vanderbilt University

Key Words: baseball, scatterplot

Scores from baseball games are bivariate data, with each variate being one team's score. Information can be obtained from viewing the distribution of baseball scores from a bivariate point of view. We offer a baseball scores plot, with each game being a point in a two-dimensional grid, and an algorithm, database, and website to allow interested users access to the plot. Using data from www.retrosheet.org, the site allows users to select Major League teams for the 1871-2006 seasons; plotted are the results from the games. We allow color-conditioning of the plot. Anomalies such as Roger Clemens' hideous lack of run support in 2005, the Pirates' dismal record in one-run games before the break in 2006 and subsequent improvement in the latter half of that year, and oddities such as tie games become evident. Rationale behind the plot techniques and website use will be discussed.

70 Small-Area Estimation: Techniques and Applications ● ✪

Section on Survey Research Methods, Social Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Hierarchical Bayes Small-Area Estimates of Adult Literacy Using Unmatched Sampling and Linking Models

❖Leyla Mohadjer, Westat, 1650 Research Blvd, Rockville, MD 20850, *leylamohadjer@westat.com*; J. N. K. Rao, Carleton University; Benmei Liu, Westat; Thomas Krenzke, Westat; Wendy VanDeKerckhove, Westat

Key Words: Small area estimation, Hierarchical Bayes estimates, Monte Carlo Simulation, Item Response Theory, Generalized Variance Functions

Funded by the National Center for Education Statistics, the National Assessment of Adult Literacy (NAAL) was designed to measure the English literacy skills of adults in the U.S. based on an assessment containing a series of literacy tasks completed by sampled adults. Sufficiently precise estimates have been produced for the nation and major subdomains of interest using the NAAL data. However, policymakers and researchers/business leaders often need literacy information for states and counties but these areas do not have large enough samples to produce reliable estimates. Therefore, small area estimation techniques are used to produce estimates of literacy levels for all states and counties in the nation. This paper describes the Hierarchical Bayesian estimation techniques used to derive a single area-level linking model to produce both county and state estimates, and credible intervals.

Hierarchical Bayes Modeling of Survey-Weighted, Small-Area Proportions

❖Benmei Liu, Westat, 1650 Research Blvd, Rockville, MD 20850, *benmeiliu@westat.com*; Graham Kalton, Westat; Partha Lahiri, University of Maryland

Key Words: weighted proportions, Hierarchical Bayes modeling;, beta distribution

When a Hierarchical Bayes area level model is used to produce estimates of proportions of units with a given characteristic for small areas, it is commonly assumed that the survey weighted proportion for each sampled small area has a normal distribution with known sampling variance. However, the assumptions of known sampling variance and normality are problematic when the small area sample size is small or when the true proportion is near 0 or 1. In an effort to overcome these problems, we propose an alternative modeling of the survey weighted proportion based on the beta distribution. We compare the results obtained from this alternative modeling with those obtained from a few commonly used modeling approaches using a Monte Carlo simulation study in which samples are generated from fixed finite population using both equal probability of selection (epsem) and non-epsem sampling designs.

A Simple Computational Method for Estimating Mean Squared Prediction Error in General Small-Area Model

❖Snigdhansu Chatterjee, The University of Minnesota, School of Statistics, 313 Ford Hall 224 Church St SE, Minneapolis, MN 55455, *snigchat@yahoo.com*; Partha Lahiri, University of Maryland

Key Words: Small area, mean squared prediction error, bootstrap, GLMM

A general small area model is a hierarchal two stage model, of which special cases are mixed linear models, generalized linear mixed models and hierarchal generalized linear models. In such models, the variability of predictors (like the empirical best predictor or the empirical best linear unbiased predictor) is usually quantified with their mean squared prediction error (MSPE). Estimators for MSPE are generally not available outside some special cases. We propose a simple resampling based estimation of MSPE for any general small area model. The proposed MSPE estimator has high order accuracy and can be guaranteed to be positive.

On Comparison of Different MSE Estimators in Small-Area Estimation

❖En-Tzu Tang, University of California, Davis, 424 Russell Park Apt 6, Davis, CA 95616, *entang@gmail.com*; Jiming Jiang, University of California, Davis

Key Words: empirical best predictors, mean squared error, nested error regression, two stage estimator

Nested-error regression models are widely used for analysis of clustered data, such as in small-area estimation. These analyses usually focus on prediction of mixed effects such as small area means. An important measure of the performance of the prediction is mean squared prediction error. In this talk, I compare several different methods for estimating the mean squared predictor error by Monte Carlo simulations. I consider situations with both normal and non-normal random effects and errors. The methods being compared include the naive, Prasad-Rao, jackknife and bootstrap methods

The Cumulative Distribution Function (CDF) Adjustment Method Applied to Small-Area Estimates of Serious Psychological Distress in the National Survey on Drug Use and Health (NSDUH)

❖Jeremy Aldworth, RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709-2194, *jaldworth@rti.org*; Neeraja Sathe, RTI International; Misty Foster, RTI International

Key Words: Design-based, Small area estimation, Respondent-level adjustments

The NSDUH is an annual national survey with a sample size of approximately 67,500, and it also provides state-level estimates using small area estimation (SAE) methods. Serious psychological distress (SPD) is measured by the K6 scale in the NSDUH. In 2003, a broad array of mental health questions preceded the K6 items. In 2005, only the K6 questions were asked. In 2004, the sample was evenly split between the long and short modules used in 2003 and 2005, respectively. Results from 2004 showed large differences in SPD prevalence rates between the two modules. Since NSDUH SAE methods require two years of data to be combined, a reversible CDF adjustment method was devised so that SPD estimates from one module in 2004 could be adjusted to be similar to those from the other module, and vice versa. This allowed us to combine 2003/2004 and 2004/2005 data to obtain state-level estimates of SPD.

Expedient Small-Area Estimation via Proportional Odds Ratio Log-Linear Models

❖Lei Li, RTI International, 7008 Valerie Anne Drive, Raleigh, NC 27613, *lei@rti.org*; Paul S. Levy, RTI International

Key Words: small area estimation, sample survey, health surveillence, loglinear model, structure preserving estimation

The synthetic methods for small area estimation are appealing due to simplicity in implementation and are suitable for categorical outcomes in situations such as public health surveillance based on sample survey data where small area estimates must be produced in a timely manner. However, they do not readily accommodate the local area characteristics. The extension of the synthetic methods to the structure-preserving estimation (SPREE) methods by Purcell and Kish opened up opportunities for incorporating local area characteristics and making use of Bayesian inference. We develop estimation procedures for SPREE via proportional odds ratio log-linear models and illustrate with data from the Behavioral Risk Factor Surveillance System and Census 2000 on employment disability in North Carolina.

Small-Area Estimation for a Survey of High School Students in Iowa

❖Lu Lu, Iowa State University, CSSM, 208A Snedecor, Ames, IA 50011, *icyemma@iastate.edu*; Michael D. Larsen, Iowa State University

Key Words: Hierarchical Bayesian Analysis, Two-Way Stratification, Sparse Design, Auxiliary Information, Generalized Linear Models

Iowa's State Board of Education conducted a stratified multi-stage sample survey to study the availability of employment preparation courses and the degree to which students in Iowa's public high schools enroll in those courses. Given the budget restrictions, the design takes a small sample of PSUs from a two-way stratification of districts, which causes high variability in direct survey estimates. A hierarchical Bayesian (HB) analysis, which borrows strength across strata with similar characteristics, is suggested to increase stability. When the number of strata increases such that the design can not assure that sample is allocated to every stratum, a design which is balanced in terms of both stratification factors is necessary and a HB predictive approach is highly recommended. The involvement of relevant auxiliary variables also should improve the estimation.

71 Ancillary Data in Survey Design and Estimation ● ✪

Section on Survey Research Methods, Social Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Handling Imperfect Overlap Identification in a Dual-Frame Survey

❖ Jay Clark, Westat, 1650 Research Blvd, Rockville, MD 20850, *jayclark@westat.com*; Marianne Winglee, Westat; Benmei Liu, Westat

Key Words: Dual-frame design, Record linkage, Threshold selection, Misclassification error

The analysis of dual frame surveys requires identification of which sampled units are included on both frames. However, this identification is imperfect when only limited and nonidentical matching data is available (except in the case of units sampled from both frames). This situation is encountered in the National Incidence Study of Child Abuse and Neglect. That study employs a dual-frame design that combines a list frame of all maltreated children investigated by Child Protective Services agencies and another sample frame compiled from maltreated children reported by sources such as the police and school staff. We handled overlap determination by using a record linkage procedure that aims to minimize the misclassification rate for matches and nonmatches. This paper describes our approach and also explores the impact of misclassification in a simulation study.

Methodology for Combining a Supplemental Sample of SSA Beneficiaries with Matched Administrative and SIPP Data

❖ Todd R. Williams, Social Security Administration, 4-C-15 Operations Bldg 4700 ME, 6401 Security Boulevard, Baltimore, MD 21235, *todd.williams@ssa.gov*; Dawn E. Haines, Social Security Administration

Key Words: supplemental sample, reweighting, standard errors

Administrative data on Social Security Disability Insurance (SSDI) beneficiaries and Supplemental Security Income (SSI) recipients lack demographic and economic characteristics. This information can be obtained by matching Survey of Income and Program Participation (SIPP) public-use files to Social Security Administration (SSA) administrative records. To ensure reliable estimates for the SSDI beneficiary and SSI recipient populations, a supplemental sample of beneficiaries is combined with the matched SIPP and SSA administrative data. Although the SIPP instrument was used to collect data for the supplemental cases, the cases are drawn using a different sample design. This paper describes the methods developed to combine these data sources, including reweighting using iterative proportional fitting. Some discussion on estimating standard errors is also given.

Using the Statistics of Income Division's Sample Data To Reduce Measurement and Processing Errors in Small-Area Estimates Produced from Administrative Tax Records

❖ Kimberly A. Henry, Internal Revenue Service, PO Box 2608, Washington, DC 20013, *Kimberly.A.Henry@irs.gov*; Robin Fisher, U.S. Department of the Treasury; Partha Lahiri, University of Maryland

Key Words: Survey Sampling, Empirical Bayes Estimation, Variance Smoothing

The large Individual Master File constructed by the Internal Revenue Service (IRS) has been used in the past to produce various income-related statistics for small geographic areas. Previous research using the Statistics of Income Division's (SOI) Form 1040 sample, a large national sample of cleaned administrative tax records, suggests the IRS data are subject to various measurement and processing errors. Thus, small-area estimates based on the IRS data, though free from the usual sampling error problem typical in small area estimation, are subject to various nonsampling errors. The SOI sample can be potentially used to reduce nonsampling errors in the IRS-based small area estimates. We propose an empirical best prediction (EBP) method to improve the IRS-based small area estimates by exploiting complementary strengths of IRS and SOI data.

An Empirical Evaluation of Various Direct, Synthetic, and Traditional Composite Small-Area Estimators

❖ William Chen, Internal Revenue Service, PO Box 2608, Washington, DC 20013, *william.w.chen@irs.gov*; Kimberly A. Henry, Internal Revenue Service; Michael Strudler, Internal Revenue Service

Key Words: General Regression Estimator, Indirect Estimators, Survey Sampling

Currently, the Statistics of Income (SOI) Division of the Internal Revenue Service uses the Individual Masterfile (IMF), administrative data for the population of Form 1040 tax returns, to produce totals of various tax return variables at the state level. Previous research based on the SOI's Form 1040 sample, a large national sample of cleaned administrative tax records, suggests that the IRS data is subject to various kinds of errors, which has lead to alternative approaches. This paper compares alternative direct, synthetic, and traditional composite estimators of state-level estimates of several variables' totals and evaluates the alternatives using various robust criteria.

Using Administrative Records To Avoid Survey Data Capture Errors

❖ Tim Withum, Lockheed Martin Corporation, 7833 Walker Dr, Greenbelt, MD 20770, *timothy.withum@lmco.com*

Key Words: Census, Administrative Records, Data Capture

The standard approach to data collection and processing generally provides the majority of the analysis as a post data capture process. This can often result in the need for the accommodation of bias that is simply due to inaccuracies that may be involved in the process of retrieving the data from a paper form or via other manually intensive processes. This paper will discuss various ways that the use of records from sources other than the primary survey can be used to help avoid such biases. More specifically, examples provided focus on the data capture processes generally used during Census processing and various methods of using administrative records to increase the processing accuracy and efficiency. In general, the approaches can be thought of as attempts to move the data analysis forward in the capture process in order to detect and correct potential biases before they occur.

Optimal Estimators for Two-Phase Sample Designs

❖ Stephen Ash, U.S. Census Bureau, 5405 8th road south, Arlington, VA 22204, *stephen.eliot.ash@census.gov*

Key Words: auxiliary information, generalized regression estimator, model-assisted

This paper extends the general result of the optimal estimator (Montanari 1987) to the two-phase sample design. We do this for several different combinations of auxiliary variables that can be available with the two-phase sample design. For each combination of auxiliary variables, we define an appropriate model and associated generalized regression estimator (Cassel et al. 1976). We then find parameters which minimize the survey variances, i.e. which make the estimators "optimal." The paper concludes with a example that compares the optimal estimators with their calibration counterparts.

72 Advances and Applications in Causal Inference and Multi-Level Modeling ● ✪

Social Statistics Section

Sunday, July 29, 4:00 pm–5:50 pm

Alternative Balance Metrics for Bias Reduction in Matching Methods for Causal Inference

❖ Jasjeet Sekhon, University of California, Berkeley, 2538 Channing Way, Survey Research Center, Berkeley, CA 94708, *sekhon@berkeley.edu*

Key Words: Matching, Causal Inference, Propensity Score, Evolutionary Program, Genetic Algorithm

Sekhon (2005; 2004) and Diamond and Sekhon (2005) propose a matching method, Genetic Matching, which algorithmically maximizes the balance of covariates between treatment and control observations via a genetic search algorithm (Sekhon and Mebane 1998). Cumulative probability distribution functions of standardized statistics are used as balance metrics. Formal hypothesis tests of balance should not be conducted as is common in the matching literature because no measure of balance is a monotonic function of bias and because balance should be optimized without limit. However, descriptive measures of discrepancy ignore information related to bias which is captured by probability distribution functions of standardized statistics. Genetic matching is able to achieve covariate balance and recover experimental benchmarks using observational data when traditional matching methods fail.

Causal Inference in Spatial Settings: A Case Study of Community Policing Programs in Chicago

❖ Natalya Verbitsky, University of Michigan, 5541 S Everett Ave, apt 310, Chicago, IL 60637, *natalya@umich.edu*; Stephen Raudenbush, The University of Chicago

Key Words: causal framework, violation of SUTVA, neighborhood research, Chicago alternative policing strategy (CAPS)

Community policing has become a popular program around the world. In contrast to its popularity, little work has been done to rigorously evaluate its effectiveness. In this paper, we examine four policy-relevant causal questions regarding the effect of community policing on Chicago's personal crime rates. To answer these questions, we extend Rubin's causal framework, which defines causal effects as unit-specific differences between potential outcomes under two alternative treatments. By employing Cox's "no interference between different units" assumption, Rubin's framework ensures that each unit has only one potential outcome per treatment. However, this assumption is not valid in spatial settings, where the treatment assignment of one's neighbors may affect one's outcome. We redefine potential outcomes and causal effects, and discuss assumptions that make this framework tractable.

Can Results from an Observational Study Be Generalized to a Larger Population? A Study of Pathways from Divorce to Illness

❖ Bridget Lavelle, Iowa State University, 204 Sndecor, Ames, IA 50011, *lavelle@iastate.edu*; Fred Lorenz, Iowa State University; Jean Opsomer, Iowa State University

Key Words: weighting, divorce

Observational studies of social phenomena often use non-randomized samples, so it is unclear whether their results can be generalized to a larger population. We build upon Lorenz et al. (2006) which found that divorced women in the Iowa Midlife Transitions (MT) Project reported more illness 10 years after divorce than their married counterparts. We construct a basic model to compare three potential mediating factors, including poorer work conditions, increased social isolation, and poorer health care, and evaluate the sensitivity of results to several hypothetical sampling scenarios. Under these scenarios, we evaluate alternative weighting and estimation schemes for the data, and assess the robustness of the findings. We argue that results which hold across multiple weighting schemes are likely to be valid more generally, despite the initial non-random selection of study respondents.

Students, Teachers, Schools, and ACT Scores

❖ Stephen Ponisciak, Consortium on Chicago School Research, 1313 E 60th St, Chicago, IL 60637, *stevep@ccsr.uchicago.edu*

Key Words: Hierarchical models, student achievement, education

In many Chicago public high schools, less than 50 percent of students made the expected gain from Plan (a "practice ACT" in fall 2003) to ACT (in spring 2005). We use a three-level hierarchical linear model to explain the differences between students' and schools' ACT scores. The most important student-level predictor of ACT performance is their prior Plan test score, while GPA and demographics are also relevant. The academic background of the teacher is important, but the performance of a student's peers is more important. Changes in the learning environment (measured by student and teacher surveys) are also linked with changes in student learning gains. When choosing between being a low-achieving student in a high-achieving school and a high-achieving student in a low-achieving school, the former seems to be the better choice, as these schools show larger gains for all students.

How Well Do Postsecondary Enrichment Programs Prepare High School Students for College? A Multilevel Models Approach

❖ Mack Shelley, Iowa State University, 323 Snedecor Hall, Department of Statistics, Ames, IA 50011-1210, *mshelley@iastate.edu*; Florence Hamrick, Iowa State University; Robyn Johnson, Iowa State University

Key Words: Multi-level models, Multivariate linear models, Education policy, Student achievement, NELS, Secondary students

We explore cumulative high school achievements in math, English, and science courses by participants in Upward Bound, Talent Search, or similar programs (UBTS) for secondary students. NELS:88 (base year through second follow-up) data were used to identify program participants as well as a comparison group of students with comparable background characteristics who did not report participation in these postsecondary enrichment programs. UBTS participation is associated with higher HS grades and accumulating more English, math, and science units in high school. Multivariate linear models and multi-level analyses demonstrate the impacts of additional student/family characteristics (at Level 1) and school-based characteristics such as racial/ethnic minority enrollment and rates of participation in free or reduced-price school meal programs (at Level 2). Key policy implications are discussed.

Multilevel Mixture Modeling Applications

Tihomir Asparouhov, Muthen & Muthen; Bengt Muthen, University of California, Los Angeles; ❖Shaunna Clark, University of California, Los Angeles, 2023 Moore Hall, Los Angeles, CA 90095-1521, *shclark@ucla.edu*

Key Words: Mixture Models, Multilevel Models, Latent Class Models, Non-Parametric Random Effects, Mplus

New developments are discussed for the analysis of multilevel data where the latent classes appear not only as level-1 (individual-level) variables but also as cluster-level variables. With categorical and count outcomes, cluster-level latent class variables also simplify maximum-likelihood computations and avoid normality assumptions for random effects by a non-parametric representation of the random effects. Several applications will be presented using the latest version of the Mplus program. Examples include multilevel analysis of achievement data with classification of both students and schools, and longitudinal analysis of mental health data using non-parametric random effects modeling. Results from standard multilevel regression and growth analysis will be contrasted with the new techniques.

Analyzing the Changes in Repeated Measures Responses in Financial Literacy Education

❖Xin Shi, University of Manchester, 196 Littlemoor Lane, Oldham, OL4 2RH, United Kingdom, *jasonshi510@hotmail.com*; Pauline Davis, University of Manchester ; Sue Ralph, University of Manchester; Valerie Farnsworth, University of Manchester; Laura Black, University of Manchester ; Afroditi Kalambouka, University of Manchester

Key Words: probabilistic, financial literacy education, logistic regression, Monte Carlo

This paper develops a general method for measuring the changes in repeated measures responses over time for single items. Data is taken from the first two cohorts of a three-year longitudinal study of financial literacy education. The logistic regression model is used to estimate changing patterns in students' personal financial management awareness and practices. Monte Carlo simulation method is applied to analyze the changes in students' personal financial management awareness and practices. Such a statistical analysis of "the changes in responses" for single items is appropriate when access to latent dimensions such as attitudinal measures is not feasible. The approach has wide application in a number of educational research contexts.

73 ASA Council of Chapters Roundtable with Coffee (fee event)

Council of Chapters

Monday, July 30, 7:00 am–8:15 am

Reaching Out to the K–12 Community

❖Jerry Moreno, John Carroll University, Dept Mathematics, 20700 North Park Blvd, University Heights, OH 44118, *moreno@jcu.edu*

Key Words: outreach, k-12 education, statistics activities

Outreach to our schools is a natural and important activity for ASA Chapters. There is more and more opportunity and need for statisticians to be a viable resource for teachers and school districts as statistics continues to become an integral part of school curricula. This roundtable will share ideas and best practices, how to get started, how to continue, and where to get funding. Activities to discuss will include science fair judging, data analysis workshops, project and poster competitions, career days, classroom visitations, school projects, adopt-a-school programs and contributing to the Statistics Teacher Network. For those of you who have not been involved but want to, come and learn. For those who have been involved, come and share! Statisticians and teachers together can, and must, make a difference!

74 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education

Monday, July 30, 7:00 am–8:15 am

Successful Group Projects

❖David Zeitler, Grand Valley State University, 1 Campus Drive, 1172 Mackinac, Allendale, MI 49401, *zeitlerd@gvsu.edu*

Key Words: team building, classroom simulation, group evaluation, group activities

Join us to discuss experiences and ideas for the design, implementation, and evaluation of group projects in statistics classes. These projects provide the opportunity for higher-level learning in ways that cannot be duplicated with homework, quizzes, or exams as students take those first tentative steps beyond manipulating data toward real understanding of information. Both instructor and student viewpoints are welcome as we explore project ideas, team setup and building, project organization and tracking, and methods for evaluation.

75 Section on Statistics and the Environment Roundtables with Coffee (fee event)

Section on Statistics and the Environment

Monday, July 30, 7:00 am–8:15 am

Hierarchical Spatial Models Bridging Ecology and Public Health

❖Lance Waller, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, *lwaller@sph.emory.edu*

Key Words: Ecology, Disease, Geographic, Information, Systems, Spatial

The field of disease ecology focuses on environmental and ecologic impacts on the emergence and spread of disease and reflects an ongoing interdisciplinary interaction between biologists, ecologists, physicians, veterinarians, and public health scientists. Such studies quantify patterns in data from genetic sequencing of pathogens, host behaviors and interactions, and landscape habitat features This goal requires flexible statistical methods that incorporated the full range of data, modeling at a variety of scales, and allowing variable data quality and accuracy. Hierarchical models provide a framework for bridging dynamic systems with multiple levels of data. Established successes appear in the fields of climate modeling and the ecology of invasive species. The goal of this roundtable is to extend such approaches to the field of disease ecology.

What's the Future for Point Process Modeling in Ecology?

❖Philip Dixon, Iowa State University, 120 Snedecor Hall, Ames, IA 50011-1210, *pdixon@iastate.edu*

Key Words: spatial pattern, ecological statistics

The analysis of spatial point patterns has a long history in ecology. Over time, ecological questions and statistical methods have changed. The threefold classification into random, clustered, or segregated patterns has been replaced by multi-scale summaries using Ripley's K function or pair correlation functions. This roundtable will gather together folks to discuss "what's next?" What questions do ecologists want to answer that can't be answered with current methods? What statistical methods might provide better answers to current questions?

76 Section on Statistics in Epidemiology Roundtable with Coffee (fee event)

Section on Statistics in Epidemiology

Monday, July 30, 7:00 am–8:15 am

Building and Using Disease Prediction Models in the Real World

❖Heejung Bang, Cornell University, 411 E. 69th St., New York, NY 10021, *heb2013@med.cornell.edu*

Key Words: risk score, risk equation, disease prediction, diabetes, kidney disease

Building risk score or prediction models is an area of mathematical/statistical modeling that can have a huge impact on patient care in real life. Accuracy and user-friendliness are two required criteria in the development of any prediction model to be not only accepted by clinicians and policymakers, but also commonly used by the general public. We will discuss popular statistical methodologies and the risk models utilized in a variety of real settings. The larger question is how do we get our "good" models into common practice? To illustrate this, I will present models for diabetes and kidney disease. Other participants may want to discuss their own prediction models or schemes and how they are being used in practice.

77 Section on Statistics in Sports Roundtable with Coffee (fee event)

Section on Statistics in Sports

Monday, July 30, 7:00 am–8:15 am

Moneyball and the NFL: Can It work?

❖Keith Schleicher, Capital One, 11813 Alder Ridge Place, Glen Allen, VA 23059, *keith.schleicher@capitalone.com*

Key Words: Sports, Football, Moneyball, NFL, Quarterback

Michael Lewis's best-seller Moneyball showed how one baseball team used statistical analysis to gain a competitive advantage. Could the same hold true for professional football? Can an NFL team help itself by using statistical analysis to inform personnel decisions, game preparation, and in-game decisionmaking? Some possible questions to consider are: 1. Can a college quarterback's performance be normalized for the game situations in which he played and the caliber of the opponent?; 2. How good is the NFL QB rating at comparing the performance of quarterbacks?; 3. If one kick can decide the outcome of a game, why is the kicker among the lower paid positions in football?

78 Section on Survey Research Methods Statistics Roundtable with Coffee (fee event)

Section on Survey Research Methods

Monday, July 30, 7:00 am–8:15 am

Web Panels: The Future of Survey Research?

❖Karol Krotki, RTI International, 701 13th St NW, Suite 750, Washington, DC 20005, *kkrotki@rti.org*

In light of current problems with traditional telephone surveys, do web surveys represent a viable option? Even if the sample is not based on a probability design?

79 Section on Teaching Statistics in the Health Sciences Roundtable with Coffee (fee event)

Section on Teaching Statistics in the Health Sciences

Monday, July 30, 7:00 am–8:15 am

Teaching Community Collaborators To Understand Health Research Methods and Results

❖Katrina Ramsey, Northwest Portland Area Indian Health Board, 527 SW Hall Blvd, Suite 300, Portland, OR 97201, *kramsey@npaihb.org*

Key Words: community collaborators, health research, teaching statistics, participatory research, communication

Many population-based studies conducted by investigators at health research agencies and academic centers involve collaboration with community-based partners. A widely used framework in tribal health projects is Community Based Participatory Research (CPBR), which involves community members in "every facet of the research process." If community members are to have substantial input into research design, analysis, and interpretation, they need some understanding of these concepts. However, lack of formal statistics training, as well as many other factors, may make it difficult for some of our collaborators to understand, and ultimately use, the information gained from studies conducted in their communities. The goal of this roundtable is to discuss successful (and possibly unsuccessful) methods for teaching and communicating with our community-based research partners.

80 Introductory Overview Lecture 2: Statistical Machine Learning and Its Applications in Bioinformatics

The ASA, IMS, ENAR, WNAR, SSC, Section on Nonparametric Statistics, Section on Statisticians in Defense and National Security

Monday, July 30, 8:30 am–10:20 am

Some Recent Advances in Classification: A Statistical Perspective

❖Ji Zhu, University of Michigan, 439 West Hall, 1085 South University Ave., Ann Arbor, MI 48109-1107, *jizhu@umich.edu*

The support vector machine and boosting are widely used tools for predicting or classifying noisy data. The ideas were introduced by Vapnik (1995) and Freund & Schapire (1996). These two methods have attracted a lot of attention due to their great success in data modeling tasks. They are now being applied to medical diagnosis, bioinformatics and genetic modeling, chemical process control, shape, handwriting, speech and face recognition, financial modeling, and a wide range of other important practical problems. In this talk, I will give a tutorial on these two methods, from the point of view of a statistician.

Applications of Statistical Machine-Learning to Modern Biological Datasets

❖Jon D. McAuliffe, University of Pennsylvania, 400 Huntsman Hall, Statistics Department, Wharton School, Philadelphia, PA 19104, *mcjon@wharton.upenn.edu*

High-throughput experimentation is now a routine part of research in biology. The corresponding need to process and analyze large, complex biological datasets has given rise to the specialized field of bioinformatics. I will give some examples of the statistical and computational issues that arise in bioinformatics analyses, and how machine-learning methods have been used with some success to address them. First I will describe the "functional annotation" problem: how to determine what biological role is played by different parts of an organism's genome. Comparing to the genomes of related organisms can help a lot; I will explain the graphical model formalism and show how it is relevant. Then I will describe the notion of "heterogeneous data integration" using support vector machines, with an application to discriminating different classes of proteins in yeast. No background in biology will be assumed.

81 Statistical Methods for Gene Regulatory Networks ● ✪

Biometrics Section, ENAR, Section on Statistics in Epidemiology

Monday, July 30, 8:30 am–10:20 am

Bayesian Error Analysis Model for Reconstructing Transcriptional Regulatory Networks

❖Ning Sun, Yale University, 300 George Street #531, New Haven, CT 06511, *ning.sun@yale.edu*; Raymond J. Carroll, Texas A&M University; Hongyu Zhao, Yale University

Key Words: gene expression, transcriptional regulation, Markov chain Monte Carlo, Bayesian, misclassification

Transcription regulation is a fundamental biological process. In this article, we propose a Bayesian error analysis model to integrate protein-DNA binding data and gene expression data to reconstruct transcriptional regulatory networks. Transcription is modeled as a set of biochemical reactions to derive a linear system model with clear biological interpretation, and measurement errors in both protein-DNA binding data and gene expression data are explicitly considered in a Bayesian hierarchical model framework. Model parameters are inferred through Markov chain Monte Carlo. The usefulness of this approach is demonstrated through its application to infer transcriptional regulatory networks in the yeast cell cycle.

Studying Coregulation and Inter-regulation of Genes via eQTL Mapping

❖Tian Zheng, Columbia University, Department of Statistics, Room 1007, 1255 Amsterdam Avenue, New York, NY 10027, *tzheng@stat.columbia.edu*

Key Words: Co-regulation, eQTL mapping, Statistical genetics

eQTL mapping is to find loci on human genome that have demonstrated linkage to or association with the expression of a gene in microarray hybridization experiments. Such identified loci may contain important information on the regulatory factors of the given gene under study. In this talk, I will discuss coregulation and inter-regulation patterns identified via similar strategies.

Self-Correcting Maps of Molecular Pathways

❖Andrey Rzhetsky, Columbia University, Dept. of Biomedical Informatics, C2B2, 1130 St. Nicolas Avenue, New York, NY 10032, *ar345@columbia.edu*; Tian Zheng, Columbia University; Chani Weinreb, Columbia University

Key Words: bayesian network, molecular network, conflicting data, text-mining

Reliable and comprehensive maps of molecular pathways are indispensable for guiding complex biomedical experiments. Such maps are typically assembled from myriads of disparate research reports and are replete with inconsistencies due to variations in experimental conditions and/or errors. It is often an intractable task to manually verify internal consistency over a large collection of experimental statements. To automate large-scale reconciliation efforts, we propose a random-arcs-and-nodes model where both nodes (tissue-specific states of biological molecules) and arcs (interactions between them) are represented with random variables. We show how to obtain a noncontradictory model of a molecular network by computing the joint distribution for arc and node variables, and then apply our methodology to a realistic network, generating a set of experimentally testable hypotheses.

Using Sequence Information to Predict Gene Regulation

❖Jun S. Liu, Harvard University, Dept. of Statistics, 1 Oxford Street, Cambridge, MA 02138, *jliu@stat.harvard.edu*; Qing Zhou, University of California, Los Angeles

Key Words: Regression trees, Gene regulation, Bayesian model average, boosting, regression, MCMC

Understanding how genes are regulated in various circumstances is a central problem in molecular biology. The adoption of large-scale biological data generation techniques such as the expression microarrays has enabled researchers to tackle the gene regulation problem in a global way. We present a method based on the Bayesian Additive Regression Trees (BART) developed by Chipman et al. (2006) for extracting sequence features to predict gene expression or enrichment values. We show that BART significantly outperformed the neural network and our earlier stepwise linear regression approaches, for real datasets including the human Oct4 and Sox2 ChIP-chip datasets and yeast amino acid starvation dataset. The variables selected by BART are also of significant biological significance.

82 Administrative Records and Data Integration ● ✪

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Monday, July 30, 8:30 am–10:20 am

Maintaining Confidentiality in Administrative and Integrated Databases

Jerome P. Reiter, Duke University; ❖ Satkartar Kinney, Duke University, *saki@stat.duke.edu*

Key Words: confidentiality, disclosure, multiple imputation

In this talk, we discuss how multiply-imputed, synthetic data can be used to maintain confidentiality when releasing microdata. We describe the synthesis of a business establishment database. The database includes many challenging variables to synthesize, including longitudinal data and skewed distributions. We focus on the analytic validity of the synthetic datasets.

Synthetic Data for Administrative Record Applications at LEHD

❖ Jeremy Wu, U.S. Census Bureau, 4700 Silver Hill Road, Suitland, MD, *Jeremy.S.Wu@census.gov*; John Abowd, Cornell University

Key Words: synthetic data, statistical validity, administrative records

The Longitudinal Employer-Household Dynamics Program at the U.S. Census Bureau has developed several synthetic data products including On The Map and a partially synthetic version of the Survey of Income and Program Participation linked to SSA and IRS data. In this paper we propose quality assurance standards that statistical agencies might use for developing and supporting such products. Documentation of inference validity and its relation to the synthetic data method of confidentiality protection is an important feature of our proposal.

Synthetic Data Disclosure Control for American Community Survey Group Quarters

❖ Rolando Rodríguez, U.S. Census Bureau, 4700 Silver Hill Rd., Stop 9100, Washington, DC 20233, *rolando.a.rodriguez@census.gov*; Robert H. Creecy, U.S. Census Bureau

Key Words: disclosure, synthetic data, group quarters, American Community Survey, ACS, microdata

Last year, we reported on an effort to construct a disclosure control method using partially-synthetic data for American Community Survey (ACS) group quarters. This effort was in anticipation of the first planned release of ACS group quarters public-use microdata samples. As ACS data for 2006 becomes available, we have been able to test our method for the first time on a full-size group quarters sample. We give results of our test, along with discussions of new modeling methods, issues faced in the synthesis workflow, and possibilities for future research.

83 Recent Advances in Survival Analysis Beyond Cox Model

IMS, ENAR, Section on Statistics in Epidemiology, Section on Risk Analysis, Biometrics Section, Biopharmaceutical Section, Section on Physical and Engineering Sciences, WNAR

Monday, July 30, 8:30 am–10:20 am

Nonproportional Hazards Models for Censored Data

❖ Zhezhen Jin, Columbia University, Dept. of Biostatistics, 722 W 168th St, New York, NY 10032, *zj7@columbia.edu*

Key Words: accelerated failure time model, general transformation model, lenear ar transformation model

In this talk, we will review three types of nonproportional hazards models for censored data: the accelerated failure time (AFT) model, linear transformation models, and general transformation models which are distribution free and with unknown monotone transformations. In particular, two general R/Splus programs for AFT model fit will be illustrated.

Semiparametric Accelerated Failure Time Model with Missing Data

❖ Menggang Yu, Indiana University Purdue University Indianapolis, Department of Medicine/Biostatistics, School of Medicine, 535 Barnhill Drive, RT 380F, Indianapolis, IN 46202, *meyu@iupui.edu*

Key Words: Accelerated Failure Time Model, Two-Phase Sampling, Missing Data, Survival Analysis, Semiparametric Method

Semiparametric accelerated failure time model (SAFT) is a popular alternative to the Cox proportional hazard model. In SAFT, log-transformed survival time is directly modeled as a sum of linear combination of covariates and an unspecified error term. Theoretical and computational properties of SAFT have been studied in the last decade when all covariates are fully observed. In this talk, we consider the case when some of the covariates may be subject to missing. We will pay special attention to a two-phase sampling scheme when some covariates are only observed in the 2nd phase of a study depending on the first phase data. Several estimation methods will be discussed and compared. Computational aspects will also be addressed.

Rank Regression, Accelerated Failure Time Model, and Empirical Likelihood

❖ Mai Zhou, University of Kentucky, Department of Statistics, Lexington, KY 40506-0027, *mai@ms.uky.edu*

Key Words: Survival Analysis, Empirical Likelihood, Censored Data, AFT model, Rank estimation, R Software package

Accelerated Failure Time (AFT) model has long been recognized as a competent alternative to the Cox proportional hazards regression model. Yet AFT models did not see much action in practical problems, partly due to the availability of (or lack of) reliable inference method and software. Recent advances in the empirical likelihood (Owen 2001) methodology and R software have provided AFT model with some nice solutions. We first discuss some recent results of empirical likelihood analysis pertaining to the rank regression of the AFT model, and then examples of computation using R software will be demonstrated.

Varying Coefficient Cox Model with Nonparametric Longitudinal Covariates

❖Jimin Ding, Washington University in St. Louis, One Brookings Dr. Cupple I, Math Department, Saint Louis, MO 63130, *jmding@math.wustl.edu*; Jane-Ling Wang, University of California, Davis

Key Words: Survival data, Longitudinal data, Cox Model, Time-varying coefficent, Joint modelling

In biomedical and clinic studies, time to some interesting events is usually recorded together with longitudinal covariates for each subject. The goal is not only to study the influence of covariates on survival time but also the pattern of longitudinal covariates. Joint modeling of the longitudinal and survival data has emerged an effective way to gain information for both processes and will be studied here. We propose to use the Cox model with time-varying coefficients to link longitudinal covariates and survival time and investigate the possible time-dependent survival regression effect. Considering longitudinal covariates are often measured intermittently and may contain measurement errors, we will estimate unobservable longitudinal process together with time-varying coefficients in the joint modeling setting.

84 Accreditation of Professional Statisticians ● ✪

Section on Statistical Consulting, SSC

Monday, July 30, 8:30 am–10:20 am

SSC Accreditation: The Canadian Model

❖Judith-Anne W. Chapman, Queen's University, NCIC Clinical Trials Group, 10 Stuart Street, Kingston, ON K7L 3N6 Canada, *JChapman@ctg.queensu.ca*

Key Words: accreditation, professional statistician

Statisticians now have formal professional recognition in UK, Australia, and Canada. The ISI is considering a similar step. Statistical Society of Canada (SSC) Accreditation of Professional Statisticians operates under recently approved Canadian Federal Trademarks, P.Stat. (Professional Statistician) and A.Stat. (Associate Statistician). There was a general opening for P.Stat. in March, 2004. With sufficient P.Stat. approved to be mentors for A.Stat., there was a first call for A.Stat. applications in June, 2006. This panelist was Chair of SSC Accreditation Committee at the time when Accreditation was opened, and will outline placement of SSC Accreditation in a multi-disciplinary and international framework.

Academia and the Professional Statistician Community

❖Mary E. Thompson, University of Waterloo, Statistics and Actuarial Science, Waterloo, ON N2L 3G1 Canada, *methomps@uwaterloo.ca*

Key Words: accreditation, professional practice, universities

Accreditation in Canada is a certification for professional practice of statistics, and has strong support from academia as well as the professional statistician community. The program has important implications for universities. It provides a new dimension to undergraduate and master's level education. University statisticians are and will be closely involved in the development and delivery of continuing education opportunities. The academic and professional statistician communities also intersect. Some reflections on the importance of this intersection and some of the questions surrounding it will be offered.

The ASA Consideration of Accreditation

❖Mary Batcher, Ernst & Young, 1225 Connecticut Ave., NW, Washington, DC 20906, *mary.batcher@ey.com*

Key Words: accreditation

Several years ago, the ASA considered certification but chose not to institute a certification program. Since then, three other statistical societies have successfully implemented accreditation programs. The ASA Task Force on Accreditation, reviewed these three programs and conducted focus groups and surveys of ASA members. Based on the interest expressed in the focus groups and survey, the Task Force recommendation to the ASA Board was the establishment of an ASA Committee to develop the basic structure of an accreditation program and present that to the Board for its consideration. This panelist led the Task Force on Accreditation and will discuss the work of the task force and its recommendations and the progress of the Committee on Accreditation.

85 Biometrics Editor Invited Session ●

WNAR, ENAR, Biometrics Section

Monday, July 30, 8:30 am–10:20 am

Multilist Population Estimation with Incomplete and Partial Stratification

❖Jason M. Sutherland, Indiana University Purdue University Indianapolis, 1050 Wishard Blvd. RG 4101, School of Medicine, Division of Biostatistics, Indianapolis, IN 46202-2872, *jmsuther@iupui.edu*

Key Words: Closed population, EM algorithm, Log-linear, Multi-list, Stratification

Multi-list capture-recapture methods are commonly used to estimate the size of elusive populations. In many situations, lists are stratified by distinguishing features, such as age or sex. Stratification has often been used to reduce biases caused by heterogeneity in the probability of list membership among members of the population; however, it is increasingly common to find lists that are structurally not active in all strata. We develop a general method to deal with cases when not all lists are active in all strata using an EM algorithm. We use a flexible log-linear modeling framework that allows for list dependencies and differential probabilities of ascertainment in each list. The method of estimating population size is applied to two examples.

Bayesian Hierarchical Spatially Correlated Functional Data Analysis with Application to Colon Carcinogenesis

❖Veera Baladandayuthapani, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd. Unit 447, Houston, TX 77030, *veera@mdanderson.org*; Bani Mallick, Texas A&M University; Raymond J. Carroll, Texas A&M University; Mee Young Hong, University of California, Los Angeles; Nancy D. Turner, Texas A&M University; Joanne R. Lupton, Texas A&M University

Key Words: Bayesian methods, Carcinogenesis, Functional data analysis, Regression splines, Semiparametric methods, Spatial correlation

We present new methods to analyze data from an experiment using rodent models to investigate the biological mechanisms surrounding p27, an important biomarker predictive of early colon carcinogenesis. The responses modeled are essentially functions nested within a two-stage hierarchy. Moreover, in our experiment, there is substantial biological motivation for the existence of spatial correlation among the functions, which arise from the locations of biological structures called colonic crypts: this possible functional correlation is a phenomenon we term crypt signaling. Thus, as a point of general methodology, we require an analysis that allows for functions to be correlated at the deepest level of the hierarchy. Analysis of this data set gives new insights into the structure of p27 expression in early colon carcinogenesis and suggests the existence of significant crypt signaling.

Bayesian Methods for Predicting Interacting Protein Pairs Using Domain Information

Inyoung Kim, Yale University; Yin Liu, Yale University; ❖ Hongyu Zhao, Yale University, 200 LEPH 60 College Street, New Haven, CT 06520, *hongyu.zhao@yale.edu*

Key Words: Bayesian method, protein interaction, domain interaction, bioinformatics, computational biology, proteomics

Protein-protein interactions play important roles in most fundamental cellular processes. Therefore, it is important to develop effective statistical approaches to predicting protein interactions based on recently available large-scale yet noisy experimental data. In this paper we propose Bayesian methods to predict protein interactions based on interactions among domains, the functional units of proteins. We also propose a new model to associate protein interaction probabilities with domain interaction probabilities. When our Bayesian methods are compared with a likelihood-based approach, our methods have smaller mean square errors through both simulations and theoretical justification under a special scenario. The large-scale protein-protein interaction data obtained from high throughput yeast two-hybrid experiments are used to demonstrate the advantages of the Bayesian approaches.

86 Exploring Models Interactively ● ✪

Section on Statistical Graphics, Section on Teaching Statistics in the Health Sciences, Section on Statistical Consulting

Monday, July 30, 8:30 am–10:20 am

Bayesian Information Analysis

❖ Aleks Jakulin, Columbia University, 1255 Amsterdam Avenue, Room 927, New York, 10027, *jakulin@stat.columbia.edu*; Andrew Gelman, Columbia University

Key Words: likelihood, information, Bayesian, Information theory

A large part of Bayesian data analysis is based on examining the posterior distributions of parameters. In comparing different models, however, parameters can change their interpretations, and it is helpful to have a more stable platform for comparison. One popular approach is to examine the posterior distribution of the likelihood, or to compare the likelihood as evaluated using different models. While the parameters of a model identify the structure, the likelihood corresponds to the ability of a model to explain the observed data. Many questions, such as ``how informative is a particular variable'' or ``what is the importance of an interaction'' are better answered in terms of a change in likelihood than in terms of parameter values. Our goal in this research is to understand any fitted complex model in terms of its simpler building blocks.

Exploratory Model Analysis: Interactive Graphical Methods for Model Selection and Comparison

❖ Simon Urbanek, AT&T Labs - Research, AT&T Shannon Lab., Room D281, 180 Park Ave, Florham Park, NJ 07932, *urbanek@research.att.com*

Key Words: exploratory model analysis, model selection, visualization, tre emethods, model ensembles

Increasing availability of computing power widens the choice of models and methods in many fields of data analysis and data mining. However, the increasing variety also makes it harder to decide which methods to use for a given problem. Many publications have been written on model selection and comparison, mainly with the goal to select "the" optimal model, usually condensing the properties into a single number. In most cases this is not adequate as models have different properties with respect to the data that cannot be measured only from one viewpoint. In this talk we discuss a different approach that attempts to understand differences in statistical models by using visualization of both data and model information. We will show a general framework for exploratory model analysis and illustrate its use on practical examples with models ranging from linear models to tree methods.

Grammatical Visualization of Statistical Models

Graham Wills, SPSS Inc.; ❖Chunling Zhang, SPSS Inc., , *clzhang@spss.com*

Key Words: statistical graphics, models

The research team at SPSS has developed an engine for the production of statistical graphics based on the principles of Wilkinson's The Grammar of Graphics. This engine, together with the specification language used to define graphics, has proved capable of producing virtually any manner of statistical graphic and also a comprehensive subset of visualizations invented in other fields. In this presentation, we will show how we have applied Grammatical principles to the visualization of statistical models, as opposed to statistical data. We will give examples of several types of statistical model and will also show an example of the visualization of a distributed model; that is, a statistical model calculated on a grid of 256 computers, each with a partition of the overall data set of one billion rows.

Exploring Models for Clustering Data

❖ Dianne Cook, Iowa State University, Dept of Statistics, 325 Snedecor Hall, Ames, IA 50011-1210, *dicook@iastate.edu*

Key Words: model-based clustering, dynamic graphics, self-organizing maps, model diagnostics, tours, EDA

Clustering data is so much fun, yet it is one of the murkiest of statistical procedures. The usual suspects measuring the goodness of fit, for example, Bayes Information Criterion for model-based clustering, often yield disappointing results. In this talk we'll describe interactive and dynamic graphics methods that help to investigate cluster analysis models. We'll focus on model-based clustering, and examine the parameterization of the variance-covariance matrix, using tours. And we'll also look at the way self-organizing maps weave their net model into multivariate data.

87 Under the Microscope: Statistical Methods for Molecular Biology ● ✪

General Methodology, Biometrics Section, Section on Statistics in Epidemiology, Section on Statistical Computing, Biopharmaceutical Section

Monday, July 30, 8:30 am–10:20 am

Efficient Design and Analysis of Genome-Wide Association Studies

❖Michael Boehnke, University of Michigan, School of Public Health, Department of Biostatistics, Ann Arbor, MI 48109-2029, *boehnke@umich.edu*

Key Words: association, genetics, type 2 diabetes, two-stage designs

Genome wide association studies in which hundreds of thousands of genetic markers are genotyped on hundreds or thousands of subjects are now a practical approach to study the genetic basis of complex human diseases. In this talk, I will discuss optimal experimental design and analysis methods for two-stage genome wide association studies in which a subset of samples is genotyped on all genetic markers in stage 1, and the remaining samples are genotyped on the most interesting markers in stage 2. I also will discuss the use of genome wide association data for matching cases and controls to correct for possible population stratification. I will illustrate these methods with application to data from stage 1 of a genome wide association study of type 2 diabetes from the Finland-United States Investigation of NIDDM Genetics (FUSION) study.

Stochastic Search Gene Suggestion: Bayesians and Biology Meet Again!

❖Michael D. Swartz, The University of Texas M.D. Anderson Cancer Center, Dept. of Epidemiology - Unit 1340, 1155 Pressler Blvd., Houston, TX 77030-3721, *mdswartz@mdanderson.org*

Key Words: Bayesian hierarchical model, Bayesian model selection, Case-control design, Gene mapping, Markov chain Monte Carlo

We describe a novel application of stochastic search variable selection to gene mapping of complex diseases (Swartz et al, Biometrics, 2006). This method introduces a dual hierarchy indicator prior that mirrors selecting alleles at loci. We also exploit the inherent hierarchical structure of SSVS applied to conditional logistic regression to incorporate linkage disequilibrium, a form of genetic correlation, into the model. We show that incorporating both the dual hierarchy indicators and the genetic dependence structure reduces the false positives and false negatives over conditional logistic regression alone. We show the methods performance on real and simulated data. Our software is available at *www.epigenetic.org/Linkage/ssgs-public.*

Statistical Models of Global Transcription Regulatory Mechanisms in Cancer

❖Wei Li, Harvard School of Public Health, 375 Longwood Avenue, LW642, Boston, MA 02115, *wli@research.dfci.harvard.edu*; ❖Xiaole S. Liu, Harvard School of Public Health, 375 Longwood Avenue, LW642, Boston, MA 02115, *wli@research.dfci.harvard.edu*

Key Words: genome tiling microarrays, model-based analysis, transcription regulation, cistrome, ChIP-chip

Chromatin Immunoprecipitation coupled with DNA microarray analysis (ChIP-chip) is a popular technique to study cistromes, the set of cis-acting targets of a trans-acting factor on a genome scale. We developed a series of algorithms for the analysis of ChIP-chip on genome tiling microarrays, including a Microarray Blob Remover (MBR) to filter probes in blob defects on the array, an algorithm for eXtreme fast MApping of Nucleotide (XMAN) probes to the genome, a Model-based Analysis of Tiling arrays (MAT) that models probe baseline behavior from probe sequence and genome copy numbers, and a web Cis-Element Annotation System (CEAS) for a comprehensive annotation of protein-bound ChIP-regions in the genome. We applied these algorithms to the ChIP-chip data of Estrogen Receptor (ER) and Androgen Receptor (AR) on Affymetrix whole human genome tiling arrays to identify their cistromes.

Scale-Based Methods in the Analysis of Proteomic Data

❖Timothy Randolph, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., M2-B500, Seattle, WA 98109, *trandolp@fhcrc.org*

Key Words: discrimination, proteomic data, multi-scale, regularization

This talk considers a variety of data from proteomic platforms and questions related to the goal of discriminating between groups. These data (noisy curves or images) are high dimensional in the sense that each datum consists of 1000s of measurements, but they are potentially (relatively) low dimensional in that they are highly locally correlated. We describe efforts to exploit local correlation and scale-based structure for normalization, feature extraction and/or regularization with the goal of class discrimination.

88

Recent Development in Econometric Time Series ● ✪

Business and Economics Statistics Section

Monday, July 30, 8:30 am–10:20 am

Cointegration Analysis with Mixed Frequency Data

❖Byeongchan Seong, Chung-Ang University, Department of Statistics, 221, Heukseok-dong, Dongjak-gu, Seoul, 156-756 Korea, *bcseong@postech.ac.kr*; Sung K. Ahn, Washington State University; Peter Zadrozny, Bureau of Labor Statistics

Key Words: missing data, Kalman filter, expectation maximization algorithm, forecasting, error correction model, smoothing

We develop a method for directly modeling cointegrated multivariate time series that are observed in mixed frequencies. We regard lower-frequency data as regularly (or irregularly) missing and treat them with higher-frequency data by adopting a state-space model. This enables us to estimate parameters including cointegrating vectors and the missing observations of low-frequency data and to construct forecasts for future values. For the maximum likelihood estimation of the parameters in the model, we use an expectation maximization algorithm based on the state-space representation of the error correction model. The statistical efficiency of the developed method is investigated through a Monte Carlo study. We apply the method to a mixed-frequency data set that consists of the quarterly real gross domestic product and the monthly consumer price index.

Studying Interactions Without Multivariate Modeling

❖Alain Hecq, University of Maastricht, Dept. of Quantitative E, P.O.Box 616, Maastricht, International 6200 MD Netherlands, *a.hecq@ke.unimaas.nl*; Gianluca Cubadda, University of Tor Vergata; Franz Palm, University of Maastricht

Key Words: Common cycles, ARIMA, Cointegration, Panel data, VAR

We propose in this paper an approach that aims at providing guidance for checking the data admissibility of non-stationary multivariate time series models (VAR or VARMA) and their implied individual ARIMA specifications. In particular we show that the presence of different kinds of common cyclical features restrictions, leading to reduced rank in the short run dynamics, explains to a large extent why we can identify such parsimonious univariate ARIMA models in applied research, a paradox that the profession had problems to explain before. This allows us to develop a new strategy for studying interactions between variables without modeling these relationships in a multivariate setting. Indeed, we develop tools to study features of individual time series with the aim to infer features of the complete system, as individual series keep a print of the system as a whole.

Bayesian Change Point Model for Time Series

❖Sinsup Cho, Seoul National University, San 56-1, Shillim-Dong, Gwanak-Gu, Seoul, International 151-742 Korea, *sinsup@snu.ac.kr*; Juwon Kim, Seoul National University; Seungmin Nam, Samsung Fire & Marine Insurance Co., LTD

Key Words: Bayesian, change point, MCMC, time series, ARCH, long memory

Change point detection in time series by Bayesian method is presented. The occurrence of the change point is modeled as the discrete time Markov process with unknown transition probabilities and is estimated by MCMC based on Chib's (1998) approach. The model assumes that all or part of the parameters in the change point model change over time and the time of the change points are known. We apply the algorithm to ARCH and ARFIMA models. Simulation is performed using a variant of perfect sampling algorithm to achieve the accuracy and efficiency. We compare the performance of the proposed change point model with the Kokoszka and Leipus (2000) CUSUM type estimator using AR(1)example. The yearly Nile-river data is analyzed as an example of the long memory process.

A Unifying Framework for Analyzing Common Cyclical Features in Cointegrated Time Series

❖ Gianluca Cubadda, University of Tor Vergata, Via Columbia, 2, Roma, 00133 Italy, *gianluca.cubadda@uniroma2.it*

Key Words: Common Cyclical Features, Reduced Rank Regression, Cointegration

This paper provides a unifying framework in which the coexistence of different form of common cyclical features can be tested and imposed to a cointegrated VAR model. This goal is reached by introducing a new notion of common cyclical features, namely the weak form of polynomial serial correlation common features, which encompasses most of the previous ones. Statistical inference is obtained by means of reduced-rank regression, and alternative forms of common cyclical features are detected by means of tests for over-identifying restrictions on the parameters of the new model. Some iterative estimation procedures are then proposed for simultaneously modeling different forms of common features. Concepts and methods are illustrated by an empirical investigation of the US business cycle indicators.

89 Causal Effects of Individualized Treatment Rules

IMS, Biometrics Section, WNAR

Monday, July 30, 8:30 am–10:20 am

Causal Inference in Sequentially Randomized Trials Based on Realistic Individualized Treatment Rules

❖ Oliver Bembom, University of California, Berkeley, 1783 Oxford St, Berkeley, CA 94709, *bembom@berkeley.edu*

Key Words: causal inference, sequentially randomized trials, marginal structural models, experimental treatment assignment assumption, IPTW, dynamic treatment rules

We consider sequentially randomized trials in which treatments are assigned in response to observed intermediate outcomes. Patients may, for example, be randomized to the remaining treatment options if they have responded poorly to their initial treatment. Conventional marginal structural models (MSMs) based on static treatment regimens often fail in this context due to a clear violation of the Experimental Treatment Assignment assumption. We show that MSMs can be used, however, to model mean counterfactual outcomes corresponding to realistic individualized treatment rules. Such rules may, for example, start patients on a particular treatment a0 and then switch them to another treatment a1 if their intermediate outcome becomes unfavorable. We present IPTW estimators for this problem and illustrate the methodology through a simulation study.

Optimal Start of Treatment Based on Time-Dependent Covariates

❖ Judith J. Lok, Harvard School of Public Health, Building II, Room 409, Department of Biostatistics, 655 Huntington Avenue, Boston, MA 02115, *jlok@hsph.harvard.edu*; James Robins, Harvard School of Public Health; Miguel A. Hernan, Harvard School of Public Health

Key Words: optimal treatment regimes, observational studies

Using observational data, we estimate the effects of treatment regimes that start treatment once a covariate, X, drops below a certain level, x. This type of analysis is difficult to carry out using experimental data, because the number of possible values of x may be large. In addition, we estimate the optimal value of x, which maximizes the expected value of the outcome of interest within the class of treatment regimes studied in this paper. Our identifying assumption is that there are no unmeasured confounders. We illustrate our methods using the French Hospital Database on HIV. The best moment to start Highly Active AntiRetroviral Therapy (HAART) in HIV positive patients is unknown. It may be the case that withholding HAART in the beginning is beneficial, because it postpones the time patients develop drug resistance, and hence might improve the patients' long term prognosis.

Estimation of the Effect of Dynamic Treatment Regimes Under Flexible Dynamic Visit Regimes

❖ Andrea Rotnitzky, Universidad Torcuato Di Tella, S·enz Valiente 1010, Buenos Aires, Argentina, *arotnitzky@utdt.edu*; Liliana Orellana, Universidad de Buenos Aire

Key Words: Marginal Structural Model, Causal Inference, Dynamic Regime

Often in the management of chronic diseases, doctors indicate the patient when the next clinic visit should be according to medical guidelines. Patients return to the clinic earlier if they need to do so. At every clinic visit treatment decisions are made. It is thus of interest to estimate the effect of dynamic treatment regimes that are to be implemented in settings in which: i) medical guidelines are used to indicate when the next clinic visit should be and these may depend on the patient health status, ii) patients may return earlier than the indicate date and iii) treatment decisions are made at each clinic visit. In this talk we present a method, based on an extension of the Marginal Structural Model (Robins 1999, 2000), which allows estimation from observational data of the effects of dynamic treatment regimes that are to be implemented in settings in which i) ii) and iii) hold.

90 The Importance of Statistical Reasoning in Law and Legal Science

Committee on Law and Justice Statistics, Section on Nonparametric Statistics, Section on Statistical Consulting

Monday, July 30, 8:30 am–10:20 am

Doctrine, District Courts, and Docketology

❖ Alan J. Izenman, Temple University, Department of Statistics, Speakman Hall (006-00), Philadelphia, PA 19122-6083, *alan@temple.edu*; David Hoffman, Temple University

Key Words: judicial opinions, dispositions, on-line databases, case-complexity index, HLM, survival analysis

An important topic of interest to legal scholars is under what conditions will a judge in a civil case decide to write an opinion in a case, instead of merely an order. Such an "opinion or order" decision tends to be reflected in whether the decision is published in either Westlaw or Lexis. To investigate the conditions under which a judge decides to write an opinion or an order, we collected a number of predictor variables related to the status of each judicial decision in a large random sample of civil cases from each of four jurisdictions, Maryland, California (ND), Pennsylvania (ED), and New York (SD), during 2003. Certain cases were not completed at the time of data collection, and so get treated as censored data. We describe an hierarchical linear model that helps deal with the complicated aspects of this question, and we present a "case complexity index" for each case.

Two Percentiles in Court: The Zuni School District Case

❖Joseph L. Gastwirth, George Washington University, 2140 Pennsylvania Ave NW, Department of Statistics, Washington, DC 20052, *jlgast@gwu.edu*

Key Words: school funding equity, measure of disparity, calculating percentiles, statutory interpretation

In 1994, the U.S. Congress rewrote the Federal Impact Aid law. Rather than give the Dep't. of Educ. the power to establish a measure of disparity to assess the equality of school funds in a state, the law specified a formula that differed in some respects. The Department interpreted the new formula to be identical to its former one. The Department determined that New Mexico would qualify as "equalized" under its interpretation but the Zuni District asserted that under the new statute the local areas not the state deserved the $60 million. The way the lower courts discussed the calculation and the speaker's views of the statistical aspects of the case will be described. As the U.S. Supreme Court has accepted the case, its resolution will also be discussed if the decision is handed down this term.

Verdict on Verdicts: Statistical Measurement of the Accuracy of Jury Verdicts

❖Bruce D. Spencer, Northwestern University, Department of Statistics, 2006 Sheridan Rd., Evanston, IL 60208-4070, *bspencer@northwestern.edu*

Key Words: statistics and the law, jury accuracy, false conviction, Type I error, Type II error, latent class

Accuracy of jury verdicts is a societal concern, always paramount to the defendant, that becomes acute during high-profile litigation. But aside from exceptional cases, little is known about how often juries as a whole get it wrong. Yet, average accuracy of jury verdicts can be studied empirically and systematically when a replication consisting of a second rating of the verdict is available. Data from a recent study of U.S. criminal cases by the National Center for State Courts (NCSC) showed only fair to poor judge-jury agreement. Simple estimates developed from the judge-jury agreement rate provide (under plausible conditions) optimistic estimates of accuracy of jury verdicts. Under stronger conditions, log-linear models provide surprising estimates of conditional and unconditional probabilities of incorrect convictions and acquittals. Further data collection is recommended.

The Effect of Serial Correlation on Statistical Evidence in Legal Cases

❖Yulia R. Gel, University of Waterloo, 200 University Ave. W., Waterloo, ON N2L 3G1 Canada, *ygl@math.uwaterloo.ca*; Weiwen Miao, Macalester College; Joseph L. Gastwirth, George Washington University

Key Words: non-parametrics, legal statistics, time series, goodness-of fit tests

When data occur sequentially in time, they typically have a positive serial correlation. Such a dependence among observations is known to affect the distribution of both parametric and nonparametric test statistics. The effect on the p-value of a particular test is related to the dependence structure underlying the data. We will present a study of the dependence effect on various statistical tests utilized in legal cases and possible remedial approaches. The implications of our results for a statistical analysis submitted in a securities law case and a jury discrimination case will be discussed.

91 Implementing GAISE in Our Classrooms

Section on Statistical Education

Monday, July 30, 8:30 am–10:20 am

Implementing GAISE in Our Classrooms

❖Carolyn P. Dobler, Gustavus Adolphus College, Department of Mathematics and Computer Science, 800 West College Avenue, St. Peter, MN 56082, *dobler@gac.edu*; ❖Carolyn K. Cuff, Westminster College, Department of Mathematics, 154 Hoyt Science Center, New Wilmington, PA 16172, *ccuff@westminster.edu*; ❖E. Jacquelin Dietz, Meredith College, Department of Mathematics and Computer Science, 3800 Hillsborough Street, Raleigh, NC 27607-5298, *dietzjac@meredith.edu*; ❖Gary Kader, Appalachian State University, Department of Mathematical Sciences, 121 Bodenheimer Drive Walker Hall, Boone, NC 28608-2092, *gdk@math.appstate.edu*; ❖Dexter C. Whittinghill, Rowan University, Department of Mathematics, 201 Mullica Hill Rd., Glassboro, NJ 08028, *whittinghill@rowan.edu*

Key Words: GAISE, Education, Teaching, Assessment

The American Statistical Association (ASA) funded and adopted the Guidelines for Assessment and Instruction in Statistics Education (GAISE) Project to provide recommendations for teaching statistics. Each of the panelists have selected at least one guideline and implemented a corresponding change or changes in their course. They will report on their challenges of making the change and whether they feel the change was successful. They will also give suggestions for others to implement changes based on GAISE.

92 Section on Health Policy Statistics Student Award Winners ● ✪

Section on Health Policy Statistics

Monday, July 30, 8:30 am–10:20 am

Evaluating Bias Correction in Weighted Proportional Hazards Regression

❖Qing Pan, University of Michigan, 2355 Arrowwood Trl, Ann Arbor, MI 48105, *qingpan@umich.edu*; Douglas Schaubel, University of Michigan

Key Words: Confidence bands, Inverse-selection-probability weights, Observational studies, Proportional hazards model, Selection Bias, Wald test

With possible biased sample, Pan and Schaubel recently proposed inference procedures for an inverse selection probability weighted Cox model, applicable when selection probabilities are estimated empirically. The proposed weighted estimator is consistent for the target population parameter, while the unweighted estimator converges to a modification of the true value. Similar statements apply to the cumulative hazard estimators. Although parameter estimation is consistent, computation is more intensive. Here, we propose methods for evaluating bias in the unweighted partial likelihood and Breslow-Aalen estimators. Asymptotic properties of the proposed statistics are derived. The finite-sample significance level and power are evaluated through simulation. The proposed methods are applied to national organ failure registry data to evaluate the bias in a post kidney transplant survival model.

Analysis of Longitudinal Data To Evaluate a Policy Change

❖Benjamin French, University of Washington, F600 Health Sciences Building, Campus Mail Stop 357232, Seattle, WA 98195-7232, *bcf@u.washington.edu*; Patrick J. Heagerty, University of Washington

Key Words: Generalized estimating equations, Generalized linear mixed models, Meta-analysis, Empirical Bayes

There are challenging aspects of policy change data with respect to analysis and inference that require consideration: defining comparison groups, accounting for heterogeneity in the policy effect, and modeling longitudinal correlation. We compare currently available longitudinal data analysis methods to evaluate a policy change. We also illustrate issues specific to evaluating a policy change via a case study of laws eliminating gun-use restrictions and firearm-related homicide. We obtain homicide rate ratios estimating the effect of enacting a shall-issue law that vary between 0.903 and 1.101. However, using methods that are most appropriate implies that enacting such a law is associated with a non-significant increase in firearm-related homicide. We conclude that in a policy change study it is essential to thoroughly model temporal trends and account for policy effect heterogeneity.

Testing for Trends in a Two-State Markov Model with Applications in Smoking Cessation Studies

❖Charles Minard, The University of Texas M.D. Anderson Cancer Center, 1155 Pressler Blvd, Unit 1340, Houston, TX 77030, *cgminard@mdanderson.org*; Wenyaw Chan, The University of Texas at Houston; Carol J. Etzel, The University of Texas M.D. Anderson Cancer Center; David Wetter, The University of Texas M.D. Anderson Cancer Center

Key Words: binary, longitudinal, GEE, Markov chain, smoking cessation

Intervention trials may observe participants alternating between two states over time. The generalized estimating equations (GEE) method is commonly used to analyze binary, longitudinal data in the context of independent variables. A trend may be evaluated by including an interaction term in the GEE model. However, the sequence of observations may also follow a Markov chain with stationary transition probabilities. Assuming a log-transformed trend parameter, determinants of a trend may be evaluated by maximizing the likelihood function. New methodology is presented here to test for the presence and determinants of a trend in binary, longitudinal observations. Empirical studies are evaluated, and comparisons are made with respect to the GEE approach. Practical application is made to data available from a smoking cessation study.

Is Newer Always Better? Re-evaluating the Benefits of Newer Pharmaceuticals

❖Michael Law, Harvard Medical School/Harvard Pilgrim HealthCare, 133 Brookline Ave, 6th Floor, Boston, MA 02215, *mlaw@fas.harvard.edu*; Karen Grépin, Harvard University

Key Words: Prescription Drugs, Propensity Score Matching, Drug Offsets, Health Care Costs

Whether newer pharmaceuticals justify their higher cost by "offsetting" other health spending (such as hospital services) is an important health policy question. We aimed to replicate a previous analysis which suggests the savings from newer drugs substantially outweigh their additional cost. We find the results are highly dependent on the model and dataset used: substituting either a model less sensitive to outliers or a newer data release results in the effect disappearing; substituting both causes it to reverse in direction. Further, we propose an alternative model using propensity score matching and a two-part expenditure model, which we estimate for hypertension. We find using a newer drug is associated with $179 higher annual drug costs, but the change in non-drug spending is indistinguishable from zero. Thus, we find no evidence of offsets from these medications.

Evaluating Predictive Capacity of Continuous Biomarkers

❖Ying Huang, University of Washington, Department of Biostatistics, Seattle, WA 98195, *ying@u.washington.edu*; Margaret Pepe, Fred Hutchinson Cancer Research Center; Ziding Feng, Fred Hutchinson Cancer Research Center

Key Words: Predictiveness Curve, risk prediction, ROC curve, cross-sectional studies

Statistical methodology is needed to critically evaluate biomarkers. A well-established criterion for marker evaluation is classification accuracy, often characterized by the ROC curve. However, classification is not always the goal. Oftentimes we use markers to predict risk of disease. Since the criteria for evaluating risk prediction markers are different from those for classification markers, we suggest an alternative to the ROC curve for their evaluation. We propose using the predictiveness curve to display the distribution of risk predicted by markers. This tool is valuable for health policy makers who are interested in policy analysis of screening and referral applied to the whole population. We propose an estimator for making inference about the curve and for making pointwise comparisons between curves in cross-sectional studies. We also develop estimators for subpopulations.

93 Design and Analysis of Method Transfer Studies in Nonclinical Pharmaceutical Development ● ✪

Biopharmaceutical Section

Monday, July 30, 8:30 am–10:20 am

Design and Analysis of Analytical Method Transfer Studies

❖James Schwenke, Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, PO Box 368, Ridgefield, CT 06877-0368, *jschwenk@rdg.boehringer-ingelheim.com*

Key Words: method transfer, equivalence testing, analysis of variance

An analytical method transfer study is a GLP study deigned to demonstrate interlaboratory reproducibility in the transfer of manufacturing. testing or assay processes from one laboratory to another. The typical study involves two laboratories with two analysts within each laboratory, testing samples over a two day period. An equivalence testing strategy is employed to demonstrate the similarity of response obtained by the two laboratories. Depending on the pharmaceutical product involved, different experimentation designs are necessary to accommodate the destructive sampling necessary for testing. Three designs will be discussed, representing independent sampling and variations on split and strip-plot designs. Examples of the analysis corresponding to each design will be presented. Studies involving three or more laboratories will be discussed through examples.

Design and Analysis of Method Transfer Studies in Nonclinical Pharmaceutical Development

❖Jinglin Zhong, Food and Drug Administration, CDER, 9758 pleasant gate lane, Potomac, MD 20854, *jinglin.zhong@fda.hhs.gov*; Yi Tsong, Food and Drug Administration

Key Words: concordance correlation, tolerance interval, individual equivalence

Analytical method transfer is an important part of analytical method development and maintenance. The current common practice of analytical method transfer is based on the equivalence of the means between the development laboratory and the production laboratory. However the most scientific sound approach would show the equivalency of individual sample readings between the two laboratories. In this talk, statistical approaches such as concordance correlation and tolerance interval approach to address individual equivalence between laboratories will be discussed. These approaches will be compared with each other and also with the mean equivalence approach on their statistical properties. Examples and simulation results will be presented to illustrate each analysis approach and the comparisons.

Some Useful Experimental Designs in Method Transfer Experiments

❖Jyh-Ming Shoung, J&JPRD, 9 Saratoga Court, East Brunswick, NJ 08816, *jshoung@prdus.jnj.com*; Stan Altan, J&JPRD

Method Transfer is a part of pharmaceutical development in which an analytical (chemical) procedure developed in one laboratory (typically the research laboratory) is about to be adopted by one or more other laboratories (production or commercial operations). The objective is to show that the recipient laboratory is capable of performing the procedure in an acceptable manner. Acceptability is typically decided on the basis of an equivalence test. But in the course of carrying out a method transfer, other questions may arise related to fixed or random factors of interest, such as analyst, apparatus, batch, supplier of analytical reagents etc. Estimates of reproducibility and repeatability may also be of interest in the course of carrying out the method transfer. This presentation will focus on the application of various block designs.

Total Error-Based Criterion for Analytical Method Transfer Analyses

❖Eric Rozet, University of Liege, Av de l Hopital 1 Bat B36, Liege, 4000 Belgium, *eric.rozet@ulg.ac.be*; Walthère Dewe, GlaxoSmithKline

Key Words: Method transfer, total error, beta-expectation tolerance interval

The objective of an analytical transfer like for method validation is to give guarantees that a laboratory masters the method of interest. As using total error has been proposed to validate a method, total error-based approaches should be also considered in transferring a method. The major difference in transfer is that only samples from a production lot are used and the sender gives, with some uncertainty, the reference value. Consequently the sender lab penalizes the acceptance limits depending on the precision of the true value estimate. Two approaches are proposed in order to have a total error-based criterion and to take into account the precision of the reference value. Firstly a beta-expectation tolerance interval is calculated and compared to the acceptance limits obtained in the sender lab. Secondly the probability to have a result outside the acceptance limits is estimated.

Approach To Assess the Equivalence Range for an Analytical Methods Transfer

❖Rene Kubiak, Boehringer Ingelheim, Dept Medical Data Services, Bingen, 55411 Germany, *rene.kubiak@ing.boehringer-ingelheim.com*

Key Words: Equivalence Range, Schuirmann, TOST, Delta Setting, OC curve, Testing specification

One advantage to use the Two-One-Sided-Tests Procedure according to Schuirmann for the assessment of the similarity of two laboratories in an analytical methods transfer is that irrelevant differences are accepted to be negligible. Since the irrelevant difference (or equivalence range) has to be defined before the transfer testing starts, it is always difficult to determine. This talk presents an approach to derive an equivalence range by taking testing specification limits, production variability, and acceptance probabilities for later (e.g., release testing procedures into consideration). Thus, statisticians can support a study director to define an equivalence range in order to control the future risk of rejecting a good quality batch.

94 Practical Considerations for Adjustment of Baseline Covariate and Center Effects in Multicenter Clinical Trials ● ✪

Biopharmaceutical Section, ENAR, WNAR

Monday, July 30, 8:30 am–10:20 am

Exploration of a Treatment-by-Region Interaction in a Clinical Trial Setting: Can the Interaction Be Explained by Clinical Factors?

❖Cristiana Gassmann-Mayer, Johnson & Johnson PRD, 1125 Trenton Harbourton Road, PO BOX 200, Titusville, NJ 08560, *cgassma1@prdus.jnj.com*; Rachel B. Weinstein, Johnson & Johnson PRD; Jesse A. Berlin, Johnson & Johnson PRD

Key Words: treatment-by-baseline factor interaction, subgroup analyses, clinical trial, post-hoc analyses

An important clinical question in clinical trials is the evaluation of treatment-by-baseline factor interactions. Differences in treatment effects across subgroups are often difficult to anticipate a priori and post hoc differences in treatment effect across subgroups are difficult to interpret. When significant treatment-by-factor interactions are observed, further data investigations should be undertaken to understand the nature and sources of the interaction. Exploratory analyses including statistical modeling and subgroup analyses should be performed with a clear understanding of power and interpretation issues. The controversial aspect of post-hoc exploratory analysis will be illustrated by an example in a clinical trial setting with a treatment-by-region interaction. Modeling strategies will be discussed, along with effective graphical and tabular data presentations of results.

Modeling Response, Interaction, and Subgroup Analyses: Some Practical Issues

❖Mohammad Huque, Food and Drug Administration, Div of Biometrics IV Room 6232 Bldg 22, 10903 New Hampshire Avenue, Silver Spring, MD 20993, *Mohammad.Huque@fda.hhs.gov*; Mohamed Alosh, Food and Drug Administration

Key Words: multiplicity adjustment, repeated measurements, subgroup analyses, testing for interaction

This presentation addresses some of the practical issues that arise in the assessment of treatment effect in clinical trials. In particular we consider: (i) repeated measurements analysis vs. final endpoint analysis, baseline adjusted vs. unadjusted analyses. (ii) lack of power and consequently under use of testing for interaction of treatment effect and (iii) susceptibility of subgroup efficacy findings to misinterpretation due to lack of specification of such analyses in the protocol and/or lack of adjustment for the multiple subgroup analyses which are frequently made. In addition, we present a flexible hierarchical approach for addressing multiplicity issue in protocol pre-specified subgroup analyses. Such an approach represents an improvement over the well-known traditional approaches which are usually so conservatives that significant results are rarely observed after adjustment.

Adjusting for Center in the Analysis of Multicenter Clinical Trials

❖Valerie Durkalski, Medical University of South Carolina, 135 Cannon Street, Suite 405, Charleston, SC 29403, *durkalsv@musc.edu*

Key Words: multicenter, clinical trials, adjusted analysis

A clinical trial comparing two or more treatments is commonly conducted at several clinical centers. This is done primarily to meet enrollment requirements in a timely manner as well as to provide evidence that the trial results are not dependent on the setting and can be generalized to a broader patient population. There are differences of opinion on whether center main effects and interaction effects should be included in the data analysis. Specific issues related to the inclusion of 'center' in the analysis will be discussed including consideration of including center in the primary or secondary analysis and consideration of treating center as a fixed versus random effect.

A Simulation Study on Covariate Adjusted Models

❖Xiaohong Huang, sanofi-aventis, 200 Crossing Blvd, Bridgewater, NJ 08807, *xiaohong.huang@sanofi-aventis.com*; Abdul Sankoh, sanofi-aventis

Key Words: covariate adjusted, power

Covariate adjusted analysis aims to adjust the treatment difference by taking into account of the unbalanced underlying characteristic between treatment groups. Simulation studies were performed to examine the effect of covariate adjustment. Power and alpha level will be compared between unadjusted and adjusted models.

Modeling Longitudinal Count Data with Possibility of Dropouts

❖Mohamed Alosh, Food and Drug Administration, Mail stop 5319 Building 22, 10903 New Hampshire Ave, Silver Spring, MD 20993, *mohamed.alosh@fda.hhs.gov*

Key Words: longitudinal count data, dropouts

The primary endpoint for many clinical trials is counts, which frequently used to reflect the severity of the disease. Examples of this include the number of lesions in dermatologic indications such as acne, basal-cell carcinoma or actinic keratosis or the number of daily seizers in epilepsy trials. For such trials the primary endpoint is usually evaluated at successive intervals during the course of the trial. One of the features of such trials is heterogeneity among subjects in their baseline counts and/or change in their expected counts during of the course of the trial. Dropouts in such trials might be related to the severity of the disease as expressed in counts. We consider random effect approach for modeling such data along with their dropouts.

95 Analysis of Recurrent Events Data ● ✪

ENAR, Biometrics Section, Section on Physical and Engineering Sciences, Section on Risk Analysis, WNAR, Section on Statistics in Epidemiology

Monday, July 30, 8:30 am–10:20 am

Semiparametric Analysis of Correlated Recurrent and Terminal Events

❖Yining Ye, Amgen Inc., 1120 Veterans Blvd, ASF 212, South San Francisco, CA 94080, *yye@amgen.com*; John D. Kalbfleisch, University of Michigan; Douglas Schaubel, University of Michigan

Key Words: Survival Analysis, Recurrent Event, Semi-parametric analysis

In clinical and observational studies, an event can recur on the same subject, and in a more complicated situation, there exists a terminal event (e.g. death), which stops the recurrent event process. In many instances, the terminal event is strongly correlated with the recurrent event process. I propose a semiparametric method to jointly model the recurrent and terminal event processes. The dependence is modeled by a shared gamma frailty that is included in both the recurrent event rate and terminal event hazard function. Marginal models are used to estimate the regression effects on the terminal and recurrent event processes. Different from the existing frailty approaches, we relax the Poisson assumption. Then we extend the analysis to model multiple types of recurrent events and terminal events.

Some Aspects of Recurrent Event Modeling

❖Edsel A. Pena, University of South Carolina, Department of Statistics, LeConte College Building, Columbia, SC 29208, *pena@stat.sc.edu*; Akim Adekpedjou, University of South Carolina; Jonathan Quiton, University of South Carolina

Key Words: recurrent event, goodness-of-fit, model validation, informative monitoring, outlier, influential observation

In this talk I will present recent results pertaining to the modeling and analysis of recurrent event data arising in biomedical settings, as well as in other scientific areas. Of particular interest are issues pertaining to model validation and goodness-of-fit, detection of outliers and influential observations, and the impact, such as loss of efficiency, of an informative monitoring period. Aside from the analytical results, we will demonstrate the procedures and issues through application to some real biomedical data sets.

A Novel Estimation Method in the Joint Frailty Models of Correlated Survival Data with Informative Censoring

❖Lei Liu, University of Virginia, 3181 Hospital West, P O Box 800717, Charlottesville, VA 22908-0717, *liulei@virginia.edu*; Xuelin Huang, The University of Texas M.D. Anderson Cancer Center

Key Words: Shared random effects model, Proportional hazards model, Dependent censoring, Survival analysis

Joint frailty models have been proposed for analysis of correlated survival data (e.g., recurrent or clustered failure times) in the presence of informative dropout or a dependent terminal event like death. The estimation in such models often involves MCEM algorithms, which is slow in convergence and difficult to implement, limiting the further application of such models. In this paper we apply Gauss quadrature techniques for estimation. The unspecified baseline hazard is approximated by a piecewise constant one. The resulting parametric model can be fitted easily by Gauss quadrature tools in standard software, e.g., SAS Proc NLMIXED. Simulation studies show such approximation yields satisfactory results for both normal and Gamma frailty models. Our method greatly facilitates the application of the joint frailty models in practice. Two real examples are given for illustration purpose.

General Transformation Models for Joint Analysis of Recurrent Events and Terminal Event

❖Donglin Zeng, The University of North Carolina at Chapel Hill, Department of Biostatistics, Chapel Hill, NC 27599-7420, *dzeng@bios.unc.edu*; Danyu Lin, The University of North Carolina at Chapel Hill

Key Words: Recurrent event, Terminal event, Transformation model, Joint model, Informative censoring

We propose a class of transformation models with random effects for joint modeling recurrent events and terminal event. The class of transformation models include both the proportional hazards model and the proportional odds model as special cases. The nonparametric maximum likelihood estimation method is used to derive the estimators, which are then shown to be consistent, asymptotically normal and asymptotically efficient. A simple algorithm is proposed to calculate the estimators. Simulation studies are conducted to examine the small-sample performance of the proposed method. The method is further applied to a real dataset.

Methods for Modeling the Recurrent Event Mean in the Presence of Time-Varying Covariates and Time-Dependent Effects

❖Douglas Schaubel, University of Michigan, Department of Biostatistics, 1420 Washington Hts, Ann Arbor, MI 48109-2029, *deschau@umich.edu*

Key Words: censoring, conditional rate, marginal mean, nonproportionality, survival analysis, terminating event

We consider the setting where the event of interest is recurrent and subject to being terminated. In addition, the treatment and treatment effect are time-dependent. Often in the presence of time-dependent effects, interest lies not in the instantaneous treatment effect, but in the treatment's cumulative effect. We propose semiparametric methods for comparing treatment-specific marginal recurrent event means. The proposal involves combining treatment-specific semiparametric mean function estimators; each of which combines the survival and conditional recurrent event rate function. Large-sample properties are derived and evaluated in finite samples through simulation. The proposed methods are applied to national kidney failure data to estimate the time beyond which the benefit of kidney transplantation is realized with respect to medical cost.

96 Recent Developments in Bayesian Variable Selection Methods ●✪

Section on Bayesian Statistical Science, ENAR, WNAR

Monday, July 30, 8:30 am–10:20 am

Bayesian Variable Selection in Gaussian Process for Cox Models

❖Naijun Sha, University of Texas at El Paso, 500 W University Ave, Bell Hall 203 Math Dept, El Paso, TX 79968, *nsha@utep.edu*; Marina Vannucci, Texas A&M University; Mahlet G. Tadesse, University of Pennsylvania

Key Words: Bayesian variable selection, Gaussian Process, Cox Model, Latent Variable, Censored Time, MCMC

In this paper, we investigate variable selection methods for Cox's proportional hazard model via a Gaussian process. We develop selection methods that allow for censored data. Our methods lead to simultaneously estimates of the survival function as well as to the identification of the factors that affect the survival outcome. We handle the problem of selecting a few predictors among the prohibitively vast number of variables through the introduction of a binary exclusion/inclusion latent vector. This vector is updated via an MCMC technique to identify promising models. We describe strategies for posterior inference and explore the performance of the methodology with simulated and real datasets.

Bayesian Variable Selection in Proportional Hazards Models with Frailties

❖Zhen Chen, University of Pennsylvania, 423 Guardian Drive, 625 Blockley Hall, Philadelphia, PA 19104, *chenz@mail.med.upenn.edu*

Key Words: Proportional hazards, Frailty, Variable selection, Cluster, MCMC, cancer

We apply a Bayesian approach to define prior distributions for frailty variances in a PH model consisting of mixtures of point masses at zero and inverse-gamma densities. This prior allocates probability to each of the models including the overall null hypothesis of homogeneity. Using a counting process formulation, the conditional posterior distributions of the frailties and proportional hazards regression coefficients have simple forms. Posterior computation proceeds via a data augmentation Gibbs sampling algorithm, a single run of which can be used to obtain model-averaged estimates of the population parameters and posterior model probabilities for testing hypotheses about the heterogeneity structure. The methods are illustrated using data from a lung cancer trial.

Nonparametric Bayes Local Regression and Variable Selection

❖Yeonseung Chung, The University of North Carolina at Chapel Hill, 222 Standish Dr, Chapel Hill, NC 27517, *chungy@email.unc.edu*; David B. Dunson, National Institute of Environmental Health Sciences

Flexibly characterizing the relationship between a response and multiple predictors has been a great interest in many applications. In such settings, interest focuses on estimating predictor-dependent distributions, while also identifying significant predictors globally and within a local region. To address these two problems simultaneously, we propose a new class of stick-breaking prior, called local Dirichlet process (lDP), for the collection of dependent distributions. A hierarchical variable selection mixture structure is incorporated in the base measure to allow uncertainty for the predictors to be included. A blocked Gibbs sampler stochastic search algorithm for the local Dirichlet process mixture (lDPM) is proposed for posterior computation. Theoretical properties are discussed and the methods are illustrated using simulated examples and an epidemiologic application.

A Variable Selection Method for Linear Models Using Modified Zellner's Prior

❖Arun Krishna, North Carolina State University, 4700 Westgrove Street, apt 1017, Raleigh, NC 27606, *akrishn@ncsu.edu*; Sujit Ghosh, North Carolina State University; Howard D. Bondell, North Carolina State University

Key Words: Variable Selection, Zellner's g-priors, Linear Model

A common method to solve variable selection problems in linear models is minimizing a penalized sum of squares, where most of the methods differ in the penalty function. Recent methods are mostly based on penalty functions including the norms of parameter estimates, or combination of them. Ridge Regression, LASSO and Elastic Net can be considered as popular examples of these shrinkage based methods. These problems can be considered under a Bayesian framework where the log of the prior density of the parameters act as penalty functions. A suitable modification of the popular Zellner's prior for regression coefficient is proposed and the penalty function is allowed to depend on the design matrix. Performances of the new prior (and hence new penalty function) are compared with some of the current variable selection methodologies using simulated and real data.

A Decision-Theoretic Approach to the Variable Selection Problem

❖Dhruv Sharma, North Carolina State University, 2824 Apt 201 Avent Ferry Road, Raleigh, NC 27606, *dbsharma@ncsu.edu*; Sujit Ghosh, North Carolina State University

Key Words: bayesian analysis, variable selection, decision theory

Variable selection within a canonical linear regression framework is a fundamental activity in the analysis of datasets. A common strategy for this problem has been to select a model that minimizes a penalized sum of squares criterion by a constraint optimization method. However the optimality of such a procedure has not been formally studied within a formal decision theoretic framework. This article presents a formal solution by obtaining the Bayes estimator corresponding to a loss function suitable for the problem. The solution is shown to be valid for the "large p small n case" case but would require a version of the stochastic search algorithm to compute the optimal estimator. We present simulation studies to compare the performance of our estimator with some of the popular variable selection approaches available. The proposed method is also illustrated using a real dataset.

97 Section on Bayesian Statistics Student Paper Competition: New Developments in Bayesian Nonparametrics ● ✪

Section on Bayesian Statistical Science

Monday, July 30, 8:30 am–10:20 am

Bayesian Inference for Directional Conditionally Autoregressive Models

❖ Minjung Kyung, North Carolina State University, 15 Wisina Ct, Medford, NJ 08055, *mkyung@stat.ncsu.edu*

Key Words: Anisotropy, Bayesian estimation, Conditionally autoregressive models, Lattice data, Spatial analysis

A spatial process observed over a set of irregular regions or a lattice is usually modeled using a conditionally autoregressive (CAR) model. The neighborhoods within a CAR model are generally connected considering only the inter-distances between the sub-regions. To accommodate the directionally different spatial effect, a new class of models is developed using different weights given to neighbors in different directions. For Bayesian estimators of the proposed model, the posterior distributions of the parameters are derived using conjugate and non-informative priors. Efficient MCMC sampling algorithms are provided to generate samples from the marginal posterior distribution. Simulation studies are presented to explore the performance of Bayesian estimators as compared to CAR model. Finally the method is illustrated using a dataset on the crime rates of Columbus, OH.

Pattern Search Optimization with a Treed Gaussian Process Oracle

❖ Matthew Taddy, University of California, Santa Cruz, SOE Grads UCSC, 1156 High Street, Santa Cruz, CA 95060, *taddy@soe.ucsc.edu*; Genetha A. Gray, Sandia National Laboratories; Herbert Lee, University of California, Santa Cruz; Robert Gramacy, University of Cambridge; Monica Martinez-Canales, Sandia National Laboratories

Key Words: optimization, expected improvement, convergence, partitioning, Gaussian process

This work combines pattern search optimization with a statistical emulator based on Treed Gaussian Processes (TGP) to create a new hybrid algorithm. The goal is to use the global probabilistic view provided by TGP to inform the local pattern search and derive a more intelligent optimization algorithm. We also propose ways in which the emulator can be used to gain information about the objective function, inform the algorithm stopping rules and provide a probabilistic analysis of the type of convergence. We present the algorithm, a framework for statistically informed optimization, and illustrate the work with numerical results.

Mixtures of Polya Trees for Flexible Spatial Frailty Survival Modeling

❖ Luping Zhao, The University of Minnesota, 1008 27th Ave SE Apt F, Minneapolis, MN 55414, *lupingz@biostat.umn.edu*; Timothy E. Hanson, The University of Minnesota; Brad Carlin, The University of Minnesota

Key Words: Areal data, Bayesian modeling, Breast cancer, Conditionally autoregressive (CAR) model, Log pseudo marginal likelihood (LPML), Nonparametric modeling

Mixtures of Polya trees offer a very flexible, nonparametric approach for modeling time-to-event data. Many such settings also feature spatial association that requires further sophistication, either at a point (geostatistical) or areal (lattice) level. In this paper we combine these two aspects within three competing survival models, obtaining a data analytic approach that remains computationally feasible in a fully hierarchical Bayesian framework thanks to modern Markov chain Monte Carlo methods. We illustrate the usefulness of our proposed methods with an analysis of spatially oriented breast cancer survival data from the Surveillance, Epidemiology, and End Results program of the National Cancer Institute. Our results indicate appreciable advantages for our approach over previous, competing methods that impose unrealistic parametric assumptions, ignore spatial association, or both.

The Nested Dirichlet Process

❖ Abel Rodriguez, Duke University, 222 Old Chemistry Bldg Box 90251, ISDS, Durham, NC 27708-0251, *abel@stat.duke.edu*; Alan E. Gelfand, Duke University; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Nonparametri Bayes, Dependent Dirichlet Process, Clustering, Gibbs sampler

In multicenter studies, subjects in different centers may have different outcome distributions. This article is motivated by the problem of nonparametric modeling of these distributions, borrowing information across centers while also allowing centers to be clustered. Starting with a stick-breaking representation of the Dirichlet process (DP), we replace the random atoms with random probability measures drawn from a DP. This results in a nested Dirichlet process (nDP) prior, which can be placed on the collection of distributions for the different centers, with centers drawn from the same DP component automatically clustered together. Theoretical properties are discussed, and an efficient MCMC algorithm is developed for computation. The methods are illustrated using a simulation study and an application to quality of care in U.S. hospitals.

Weak Consistency of General Bayesian Kernel Mixture in Density Estimation

❖ Yuefeng Wu, North Carolina State University, 1251 University Ct Apt 304, Raleigh, NC 27606, *ywu9@ncsu.edu*; Subhashis Ghoshal, North Carolina State University

Key Words: Posterior consistency, mixture, Bayesian, nonparametric, density estimation

The Bayesian kernel mixture models for estimating probability densities are used in different model structures and with many different kernels. Posterior consistency is not studied for all these models. In this paper, we unify these models in a general form and give conditions under which the posterior weak consistency can be achieved in such general form. It is an improvement over the existing results since it covers a wide variety of kernels. As examples, we give the conditions for posterior weak consistency in the models using normal, skewed normal, t, Laplace, Logistic, Beta, Triangular, log-normal, Gamma and Weibull densities as kernels.

98 Applications of Machine Learning and Network Tomography in Statistics ● ✪

Section on Physical and Engineering Sciences
Monday, July 30, 8:30 am–10:20 am

Making the Best Use of Available Data: The Presence-Only Problem in Ecology

❖Gillian Ward, Stanford University, Department of Statistics Sequoia Hall, 390 Serra Mall, Stanford, CA 94305, *gward@stanford.edu*; Trevor Hastie, Stanford University

Key Words: presence-only problem, positive and unlabeled examples, EM algorithm, gradient boosting, boosted trees, logistic regression

Rich resources of species presence records are becoming freely available electronically, but they typically do not include records of species absence. A similar problem arises in text categorization, that of positive and unlabeled examples. As data collection can be prohibitively expensive, we would like to use these presence-only data to estimate a presence-absence model of species distribution across a landscape. We critique Maxent, an existing model, and propose two new methods: an EM algorithm that can be used with almost any off-the-shelf logistic model, and a gradient boosting model that is available in R. Both methods require an external estimate of overall population prevalence.

An Iterative Algorithm for Extending Learners to a Semisupervised Setting

❖Mark Culp, University of Michigan, 1843 Pointe Crossing 201, Ann Arbor, MI 48105, *culpm@umich.edu*; George Michailidis, The University of Michigan

Key Words: Machine learning, Non-parametric smoothing, Additive Models, CART, Semi-Supervised

In this talk, we present an iterative algorithm, whose objective is to extend learners from a supervised setting into a semi-supervised setting. The algorithm is based on using the predicted response values for observations where it is missing (unlabeled data) and then incorporates the predictions appropriately at subsequent stages. Convergence properties of the algorithm are investigated for particular learners, such as linear/logistic regression, kernel smoothers, generalized additive models, tree partitioning methods, partial least squares, etc. The algorithm is illustrated on a number of real data sets using a varying degree of labeled response.

Evidence Contrary to the Statistical View of Boosting

❖David Mease, San Jose State University, Department of Marketing and Decision Sci, College of Business BT 551, San Jose, CA 95192-0069, *mease_d@cob.sjsu.edu*

Key Words: AdaBoost, Boosting Algorithms, LogitBoost

The statistical perspective on boosting algorithms focuses on optimization, drawing parallels with maximum likelihood estimation for logistic regression. In this talk we present empirical evidence that raises questions about this view. Although the statistical perspective provides a theoretical framework within which it is possible to derive theorems and create new algorithms for general contexts, we show that there remain many unanswered important questions. Furthermore, we provide examples that reveal crucial flaws in the many practical suggestions and new algorithms that are derived from the statistical view. We examine experiments using simple simulation models to illustrate some of these flaws and their practical consequences. This is joint work with Abraham Wyner at the University of Pennsylvania.

Adversarial Classification

❖Bowei Xi, Purdue University, 250 N. University Street, West Lafayette, IN 47907, *xbw@stat.purdue.edu*; Murat Kantarcioglu, The University of Texas at Dallas; Christopher Clifton, Purdue University

Key Words: Adversarial learning, Classification, Game theory

Many data mining applications are faced with active adversaries. In all these applications, initially successful classifiers could degrade easily. This becomes a game between the adversary and the data miner. The adversary modifies its strategy to avoid being detected by the current classifier and the data miner then updates its classifier based on the new threats. We investigate the possibility of an equilibrium in this seemingly never ending game, where neither party has an incentive to change. Modifying the classifier causes too many false positives with too little increase in true positives. Changes by the adversary decrease the utility of the false negative items that aren't detected. We develop a game theoretic framework where the equilibrium behavior of adversarial learning applications can be analyzed, and provide a solution for estimating the equilibrium point.

A Somewhat General Framework for Active Network Tomography

❖Earl Lawrence, Los Alamos National Laboratory, MS F600, Los Alamos, NM 87545, *earl@lanl.gov*

Key Words: network, tomography, trees, mixture models

The field of active network tomography is concerned with the estimation of link-level performance measures (e.g., delay distributions for packets on a link) based on measured end-to-end performance of injected traffic (e.g., total path delay for a probe packet). One area of continuing research is the choice of appropriate distributional forms for estimating delay as many standard parametric distributions are inappropriate. In particular, most parametric distributions are inadequate for modeling the tail behavior of delay distributions. This talk will explore the use of mixture modeling to overcome this limitation. Mixture models provide a flexible tool for capturing overall shape and tail behavior. We will also discuss simpler models with a semiparametric specification for moments that can be used in monitoring scenarios. This framework allows fast, simple estimation.

99 ACS Multi-Year Estimates Methodology and Results

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section, SSC
Monday, July 30, 8:30 am–10:20 am

Methodology for the Production of American Community Survey Multiyear Estimates

❖Anthony Tersine, Jr., U.S. Census Bureau, 4600 Silver Hill Rd, 4H477, Washington, DC 20233, *anthony.g.tersine.jr@census.gov*; Mark Asiala, U.S. Census Bureau

Key Words: American Community Survey, Estimation

The American Community Survey (ACS) will be producing three- and five-year period estimates starting in 2008 and 2010, respectively. Before these are released the ACS has created a series of period estimates for research using the data in thirty-four counties from data collected in 1999–2005.

These estimates were produced using the methods planned for 2008 and 2010. We will share our research results on the methods that the Census Bureau has developed to produce the multi-year estimates. There are four dimensions of the methodology that we will emphasize: the conceptualization of the estimates, the estimation process, changes in geographic definitions, and use of inflation factors for dollar-valued estimates.

Imbedding Model-Assisted Estimation into ACS Estimation

❖Robert Fay, U.S. Census Bureau, 4700 Silver Hill Rd, Washington, DC 20233, *robert.e.fay.iii@census.gov*

Key Words: American Community Survey, Calibration estimators, GREG, survey estimation

The American Community Survey (ACS) began full implementation in 2005. Estimation for ACS, as typical for other large Census Bureau surveys, uses a complex series of ratio-estimation and other adjustments to the weights. Working originally from ACS data for 1999–2001 in 36 test counties, previous research suggested that multiyear tract-level estimates could be improved by imbedding a step of model-assisted estimation, specifically generalized regression estimation (GREG), in the current ACS estimation. In particular, the GREG step incorporates administrative record data. Most data sets produced for the ACS Multiyear Analysis Project incorporate the GREG step, but alternative sets of 3- and 5-year estimates without the GREG step were produced for purposes of comparison. The paper will describe new refinements in the estimation approach and its potential future role in ACS estimation.

Analysis of Variance Estimates from American Community Survey Multiyear Estimates

❖Michael Starsinic, U.S. Census Bureau, 14607 London Lane, Bowie, MD 20715, *mds449@earthlink.net*; Anthony Tersine, Jr., U.S. Census Bureau

Key Words: American Community Survey, Variance estimation

The American Community Survey (ACS) will produce a series of three- and five-year period estimates for 34 ACS counties using data from 1999–2005. This series of period estimates will provide a large amount of information regarding the variance distributions for multi-year ACS estimates for the first time. The focus of this paper is on answering three questions: how much a new weighting methodology reduced variances at the place and census tract levels, what are the variance characteristics of estimates at the block group level (the smallest geographic level published in the ACS five-year products), and the relationship between the variances of one-, three-, and five-year period estimates for specific characteristics. We will share the results of this research.

Statistical Issues and Interpretation of the American Community Survey's One-, Three-, and Five-Year Period Estimates

❖Michael Beaghen, U.S. Census Bureau, 6641 Wakefield Drive 518, Alexandria, VA 22307, *michael.a.beaghen@census.gov*

Key Words: long form, data products, usability of data products

The Census Bureau is replacing the decennial long form with the American Community Survey (ACS). The ACS differs from the long form in important ways. While the long form represented a snapshot in time, for example, April 1, 2000, the ACS will draw monthly samples over the ten-year intercensal period. Also, instead of one estimate, the ACS will produce three estimates: one based on one year of collected data, one based on three years of collected data, and one based on five years. These changes yield novel and interesting questions of interpretation. This paper presents results of research aimed at developing guidelines to aid ACS data users in interpreting the ACS's one-year, three-year, and five-year estimates, and to help them determine which of these estimates is most appropriate for a particular application.

100 Recent Developments in Statistical Methodology for Diagnostic Medicine ● ✪

WNAR, Biometrics Section, ENAR, Section on Health Policy Statistics

Monday, July 30, 8:30 am–10:20 am

Noniterative, Semiparametric, Least-Squared Method of ROC Curve Estimation

❖Xiao-Hua (Andrew) Zhou, University of Washington, 1705 Pacific Street NE, Seattle, WA 98195-7232, *azhou@u.washington.edu*; Liansheng Tang, University of Washington

Key Words: ROC curve, Semi-parametric, Biomarkers

The receiver operating characteristic (ROC) curve provides an ideal way to evaluate the discriminating ability of diagnostic tests. In this talk, we introduce a simple estimation method for a semiparametric ROC curves, particularly clustered ROC curve data. Unlike other existing methods, the new approach does not require an iterative algorithm and is easy to implement. Moreover, a nice property of this method is the invariance of estimated parameter vector to any monotone transformation of the measurement scale. We show that the parameter vector in ROC curves is consistent under mild assumptions and derive a consistent estimator of its asymptotic covariance matrix. We derive the theoretical simultaneous confidence bands of estimated ROC curves. The finite sample performance of the proposed procedure is evaluated using Monte Carlo simulations.

Statistical Methods for Differential Diagnosis of Second Primary Cancers from Metastases

❖Colin B. Begg, Memorial Sloan-Kettering Cancer Center, 307 East 63rd St, 3rd Floor, New York, NY 10021, *beggc@mskcc.org*

Key Words: metastases, statistical, clonal, cancer, differential, tumors

In recent years numerous investigators have conducted studies of pairs of tumor specimens from the same patient to determine whether the tumors share a clonal origin. These studies can be of considerable clinical significance, especially in settings where the distinction of a new primary cancer and metastatic spread of a previous cancer would lead to radically different indications for treatment. Studies of clonality have typically involved comparison of the patterns of somatic mutations in the tumors at candidate genetic loci to see if the patterns are sufficiently similar to indicate a clonal origin. More recently, some investigators have explored the use of array CGH for this purpose. In this talk new statistical approaches to this problem will be presented that are suited to each of these types of data, along with the results of studies to examine the validity of the methods.

Design of Diagnostic Accuracy Studies for PPV and NPV

❖Jason Fine, University of Wisconsin-Madison, K6 420 Clinical Science Center, 600 Highland Avenue, Madison, WI 53792, *fine@biostat.wisc.edu*; David Steinberg, Tel-Aviv University; Richard Chappell, University of Wisconsin-Madison

Key Words: Case-control study, Diagnostic testing, Optimal allocation

Important properties of diagnostic methods are their sensitivity, specificity, and positive and negative predictive values (PPV and NPV). These methods are often assessed via case-control samples, which include one cohort of cases, known to have the disease, and a second control cohort of disease-free subjects. Such studies give direct estimates of sensitivity and specificity, but only indirect estimates of PPV and NPV, which also depend on the disease prevalence in the tested population. We develop formulas for optimal allocation of the sample between the case and control cohorts and for computing sample size when the goal of the study is to prove that the test procedure exceeds pre-stated bounds for PPV and/or NPV for pre-specified prevalence range. Surprisingly, the optimal sampling schemes for many purposes are highly unbalanced, even when information is desired on both PPV and NPV.

A Unified Approach to Nonparametric Comparison of Receiver Operating Characteristic Curves for Longitudinal and Clustered Data

Gang Li, University of California, Los Angeles; ❖Kefei Zhou, Amgen Inc., 1710 W Hillcrest Dr, Apt 96, Newbury Park, CA 91320, *kzhou@amgen.com*

Key Words: ROC curve, Nonparametric comparison, clustered data, quantile process, partial area under the curve, confidence band

This article presents a unified nonparametric method for comparisons of ROC curves with clustered data. Most previous methods are developed for a single index, and the asymptotic variance usually involves unknown densities whose nonparametric estimates are unstable. We derive the joint asymptotic distribution of correlated empirical ROC curves. We then develop a Monte Carlo method to approximate their joint distribution without involving density estimation. We prove that the Monte Carlo method is consistent. The theory is applied to derive inferential procedures for comparing sensitivities at a fixed specificity and weighted areas under ROC curves, construct confidence bands for the difference of ROC curves, estimate the set of specificities at which a diagnostic test is more sensitive than the other. Simulations showed that our methods work well for small samples.

101 Harnessing the Power of the ASA Archives: Preserving and Accessing ASA's Institutional Documentary Heritage ●

Committee on ASA Archives and Historical Materials, Committee on Committees, Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Monday, July 30, 8:30 am–10:20 am

Harnessing the Power of the ASA Archives: Preserving and Accessing the ASA's Institutional Documentary Heritage

❖Frank J. Potter, Mathematica Policy Research, Inc., PO Box 2393, Princeton, NJ 08543, *fpotter@mathematica-mpr.com*; ❖Melissa Gottwald, Iowa State University, Special Collections, 403 Parks Library, Ames, IA 50010-2140, *gottwald@iastate.edu*; ❖John P. Deley, Energy Information Administration, 1000 Independence Ave SW, Forrestal Building EI30, Washington, DC 20585, *johnpaul.deley@eia.doe.gov*; ❖Ken Williams, Utah State Archives and Records Service, 346 S. Rio Grande Street, Salt Lake City, UT 84101, *kenwilliams@utah.gov*

Key Words: archives, context, documentary heritage, history, Iowa State University, statistical theory and methods

The preservation of ASA's documentary heritage constitutes a significant investment in the Association's future. The Committee on ASA Archives and Historical Materials works closely with the Iowa State University Archives, Parks Library, to preserve, store and provide researcher access to the institutional holdings of ASA. This session will provide a virtual tour of ASA's collections and the treasures it contains. Three speakers will focus on: reasons for our archives, evolutions in statistical theory and methods, changing avenues of statistical historical expression, access strategies and policies, and methodologies for better understanding the context these vital historical assets provide contemporary decision makers and future students of our profession. If the "Past is Prologue" then ASA's Archival records hold the key to ASA's dynamic future.

102 Recent Advances in Microarray Data Analysis V ● ✪

Biometrics Section

Monday, July 30, 8:30 am–10:20 am

Microarray Gene Expression Analysis of Soybean Genome Using R-Bioconductor

❖Jean Roayaei, National Institutes of Health, Room 28, 362, Miller Drive Frederick, MD 21702, *roayaeij@css.ncifcrf.gov*

Key Words: microarray, gene expression, Affymetrix GeneChip, quality control of chips, empirical Bayes t, volcano plot

A popular application of microarray technology is the identification of genes that are differentially expressed in two different populations. With Affymetrix GeneChip technology, numerous steps are required in first, establishing the quality control of chips, and, second, in selecting interesting genes. Due to the cost involved in Affymetrix microarray technology, experiments using small numbers of arrays are common. In this research, we processed three Hawaii/Ressitent and three Taiwan/Susceptible arrays. This paper describes specific procedures for (1) conducting quality assessment of Affymetrix GeneChip soybean genome data and (2) performing analyses to determine differential gene expression using Bioconductor and R. All biostatistical analyses performed in this research were conducted using R version 2.4.0 and Bioconductor version 1.9.

Conjugate Hierarchical Modeling of the Error Variance in Tests for Differential Gene Expression

❖Jason Obsorne, North Carolina State University, 2501 Founders drive, 16 Patterson Hall, Raleigh, NC 27695-8203, *osborne@stat.ncsu.edu*

Key Words: microarray, shrinkage estimation, optimal discovery procedure, regularization

A conjugate mixing distribution for the error variance is proposed to model inhomogeneity of gene variances in microarray experiments. The likelihood obtained by integrating over this conjugate is maximized to estimate mixing parameters. The predictive distribution of the error sums of squares from linear gene models is more flexible than the usual chi-square distribution and can improve fit. The estimated conditional mean of the variance shrinks the gene-specific variances towards an estimate that pools over all genes, with weights on gene-specific error mean squares proportional to the number of replicate arrays, thereby sensibly borrowing strength across genes. Simulations are used to investigate improved sensitivity of the resulting regularized F-tests as well as performance when the shrinkage technique is used for estimation of Storey's optimal discovery procedure.

False Discovery Rate in Microarray Studies

❖Moonsu Kang, The University of North Carolina at Chapel Hill, Department of Biostatistics, School of Public Health, 3101 McGavran Greenberg Hall CB 7420, Chapel Hill, NC 27599-7420, *mkang@email.unc.edu*

Key Words: DNA microarrys, multiple hypothesis testing, False discovery rate

An important aim of DNA microarrays is the identification of differentially expressed genes leading to multiple testing. False Discovery Rate (FDR) is a new approach to multiple comparisons problems. However, current FDR procedures control the FDR under independence or unduly restrictive dependence structures, resulting in an unreliable FDR estimation. We propose two-stage FDR controlling procedure under suitable dependence structures. This proposed testing procedure is based on a Poisson distributional approximation. It eliminates the need to consider specific dependence structures of genes. We compare the performance of the proposed FDR procedure with that of other FDR controlling procedures, with illustration of Leukemia study of Golub et al and simulated data. In these studies, compared to other procedures, the proposed FDR procedure has greater power without much elevation of FDR.

Sparse Partial Least Squares Regression with an Application to the Genome Scale Transcription Factor Activity Analysis

❖Hyonho Chun, University of Wisconsin-Madison, 1300 University Avenue, Madison, WI 53706, *chun@stat.wisc.edu*; Sunduz Keles, University of Wisconsin-Madison

Key Words: SPLS, variable selection, Gene expression, genome-wide binding data

Partial least squares (PLS) has been used in analysis of modern biological data which involves high-dimensionality and multicollinearity. However, PLS is not particularly tailored for variable selection, and this could be problematic when majority of the variables are noise. We show inconsistency of PLS in the presence of large number of noise variables. We propose a sparse partial least squares (SPLS) which aims to simultaneously achieve good predictive performance and variable selection thereby producing sparse linear combinations of the original predictors. We formulate SPLS by imposing L1 penalty, and show that simple soft thresholding is the solution for univariate response. We investigate the performance of SPLS by simulation study and apply SPLS to the problem of inferring transcription factor activity by integrating gene expression microarray data and genome-wide binding data.

Conservatively Optimal Multiple Testing for Comparative Microarray Experiments

❖James Signorovitch, Harvard University, Dept. of Biostatistics, 655 Huntington Ave., Boston, MA 02115, *jsignoro@hsph.harvard.edu*

Key Words: gene expression, optimal multiple testing, empirical Bayes, false discovery rate

Optimality theory for multiple hypothesis testing is considered in the context of gene expression microarray experiments. We show that optimal testing procedures are generally not identifiable but by conditioning on certain sufficient statistics we can define near-optimal procedures that are identifiable and conservative. In simulations and in real data, estimates of these near-optimal procedures are shown to substantially outperform the estimated Optimal Discovery Procedure (Storey et al. 2005) and the Empirical Alternative Hypothesis (Signorovitch 2006) by exploiting strong patterns in the data that are generally ignored.

Improved Significance of Microarrays

❖Shunpu Zhang, University of Nebraska-Lincoln, Department of Statistics, Lincoln, NE 68583-0963, *szhang3@unl.edu*

Key Words: microarray, null statistics, test statistics, false discovery rate, false positive, fudge factor

The Significance Analysis of Microarrays (SAM) is a popular method for detecting significantly expressed genes and controlling the false discovery rate (FDR). However, it has been recently reported in the literature that SAM tends to over-estimate the FDR. In this paper, we propose an improved significance analysis of microarrays (ISAM) method. The improvement is achieved by employing the more efficient test and null statistics of Zhang (2006) and a novel way of choosing the fudge factor. Through extensive simulations, we show that our proposed method consistently outperforms SAM in the sense that it tends to contain smaller number of true false positive (FP) given that the same number of significant genes is identified by SAM and our proposed method. We also demonstrate that our proposed method provides reasonably unbiased estimates of the FDR.

Comparison of Meta-Analytical Approaches for Gene Expression Profiling

❖Heather Adams, University of Illinois at Urbana-Champaign, 1207 W Gregory Drive mc630, Urbana, IL 61801, *hadams1@uiuc.edu*; Sandra Rodriguez-Zas, University of Illinois at Urbana-Champaign; Bruce Southey, University of Illinois at Urbana-Champaign

Key Words: microarray, mixed effects models, meta-analysis

Meta-analysis allows the integration of information across studies, enhancing the estimation of effects evaluated across experiments. We evaluated two approaches to combine microarray data that have complementary advantages. Results from meta-analysis of summary results and from joint analysis of raw gene expression data pertaining to four studies profiling gene expression of honey bees at two adult maturation stages were compared. A total of 192, 89, 2207, and 5 genes had significant differential expression ($P< 0.01$) within study, 1030 and 31 genes were significant in the joint and summary meta-analyses, respectively. These comparable results suggest that although joint analysis accounting for within and between studies sources of variation is desirable, the limited availability of raw data and descriptions necessary for adequate modeling make meta-analysis a suitable alternative.

103 Analysis of Censored Data and Cox's Model ● ✪

Biometrics Section

Monday, July 30, 8:30 am–10:20 am

Minimum Distance Estimation in Two-Sample Scale Problem Under the Partial Koziol-Green Model

❖Ke Wu, California State University, Fresno, Department of Mathematics, 5245 N Backer Ave Mail Stop PB108, Fresno, CA 93740, *kewu@csufresno.edu*

Key Words: minimum distance estimation, partial Koziol-Green model, partial ACL estimator, scale estimation, asymptotic properties, random censorship

This paper considers some Cramer-Von Mises type minimum distance estimators of scale parameter in the two-sample scale problem with randomly right censored data under the partial Koziol-Green (PKG) model. In our minimum distance estimator of scale parameter, the partial ACL estimator for the survival function under the PKG model of random censorship is used. The paper proves the asymptotic normality and consistency of the minimum distance estimator. A comparison of the estimator using the partial ACL estimator of survival function and the estimator of Koul and Yang (1989) using the product limit estimator (PLE) of Kaplan and Meier (1958) for the general right censoring model is made under the PKG model.

Sensitivity of Cox Regression to Nonignorable Censoring

❖Tao Liu, Brown University, 48 Thomas Olney Common, Providence, RI 02904, *tliu@stat.brown.edu*; Daniel F. Heitjan, University of Pennsylvania

Key Words: informative censoring, ignorability, sensitivity analysis, coarse-data model, ISNI

The validity of the partial likelihood in the Cox proportional hazards model for censored survival data is contingent on the assumption that the censoring mechanism is ignorable. Nonignorable censoring can spoil inferences made under an ignorable model. In this paper, we propose to conduct a local sensitivity analysis to address this issue, assuming that the true censoring mechanism is unknown but not far from ignorable. We express the censoring mechanism by the coarse-data selection model, and derive a profile likelihood to account for the random censoring. We show that the index of local sensitivity to nonignorability (ISNI) (Troxel et al., 2004) can be seamlessly extended to the profile likelihood, and we apply it to assess local sensitivity of the Cox regression to nonignorability. We demonstrate our method using data from a study of prognostic variables in advanced lung cancer.

Dealing with Tied Observation in Estimating Survival Function

❖Jiantian Wang, Kean University, 1000 Morris Ave, Union, NJ 07083, *jwang@kean.edu*

Key Words: censored failure time, survival function, tied observation

Under random censoring model, tied observation supposedly can not happen. However, under general censoring models, or, when both failure time and censoring time are discrete, there can be a significant proportion of tied observations. In this talk, we discuss the approaches to dealing with tied observations in estimating survival function. For a noninformative model with physical background, we verify the validity of the standard failure-before-censoring method. In the case when both of failure and censoring time are discrete, we give out a necessary and sufficient condition on the identifiability of survival function and derive several corresponding estimators associated with different observing mechanisms.

Test of Fit to a Semiparametric Bivariate Family with Censored Data

❖Shulamith Gross, Baruch College/CUNY, One Baruch Way, New York, NY 10010, *shulamith_gross@baruch.cuny.edu*

Key Words: bivariate, test of fit, censoring, semiparametric

In this talk I shall introduce a test of fit to a semiparametric family of Bivariate distributions recently introduced by Gross and Huber. The general family is multivariate and is particularly appropriate for reliability applications in a multi-component system, or in dependent competing risk problems. The family was proposed as a possible alternative to the semiparametric copula models or frailty models for multivariate data, and adapts readily to applications with censored data. The test is distribution free in the uncensored case. The power and asymptotic validity of the proposed test of fit will be studied via simulations. The talk is based on joint work with Catherine Huber.

Quantile Regression for Censored Point Processes

❖Rajeshwari Sundaram, National Institute of Child Health and Human Development, 6100 Executive Boulevard, NIH, Room 7B05 BMSB DESPR NICHD, Rockville, MD 20852, *sundaramr2@mail.nih.gov*; Subhashis Ghoshal, North Carolina State University

Key Words: Recurrent Events, Censoring, Survival

Recurrent events data often arise in many longitudinal follow-up studies, for example, repeated occurrences of tumors in cancer patients. In analyzing such data, majority of the work have focused on intensity based models or models for the mean function of the underlying recurrent events process. As an alternative to such models, we propose a semiparametric regression model for the quantile of the recurrent events process. In fact in dealing with recurrent medical costs related to episodic events, it may be more desirable to model the quantiles instead of the mean as the distribution of the medical costs are typically skewed. We propose an estimating equations based approach for estimating underlying parameters of interest. Finite sample properties will be investigated through simulations. We will present their large sample properties and conclude with a real data illustration.

Multivariate Rank-Order Two-Sample Test for Censored Data: More Intuitive Expression of the Treatment Effect

❖Edward Mascha, The Cleveland Clinic, 11781 Pinewood Trail, Chesterland, OH 44026-1812, *maschae@ccf.org*

Key Words: time-to-event, distribution-free, multivariate, generalized odds ratio, Mann-Whitney difference, stochastic ordering

Multiple time-to-event tasks are sometimes used to assess health of stroke patients, for example, where interest lies in the treatment effect between two groups on the vector of responses. With no censoring, a multivariate rank sum test can assess equality of groups. For the censored case, Wei and Lachin (1984) and Thall and Lachin (1988) proposed various distribution-free tests to assess homogeneity of the treatment effect across multiple outcomes, a general test of association, and tests of stochastic ordering. Interpretation for the Wei-Lachin tests is based on the Mann-Whitney difference (MWD), the difference between the probability (X>Y) and probability (X<Y). We apply these methods to stroke data and then adapt the generalized odds ratio (GOR) of Agresti (1980) to this multivariate censored case, giving a more intuitive interpretation than the MWD for such data.

CUSUMs Based on the Cox Model To Monitor Outcomes at Multicenter Studies

❖Pinaki Biswas, University of Michigan, Dept of Biostatistics, 1420 Washington Hts, Ann Arbor, MI 48109, *pbiswas@umich.edu*; John D. Kalbfleisch, University of Michigan

Key Words: CUSUM, Wald's SPRT, Cox regression, Poisson process, Wald's Identity

In organ transplant it is important for a particular transplant center to assess the efficacy of transplants performed over time and provide a signal if there is evidence of poor performance. The CUSUM has been suggested as a sequential monitoring scheme in Page (1954) and has also been implemented in Axelrod et al. (2006) and Steiner et al. (2001). In the above articles, the unit of analysis is a discrete outcome associated with the transplant. In this paper, we utilize the CUSUM statistic based on a Cox regression model prospectively and update the event status continuously in time, also accounting for patient-mix covariates. Simulations demonstrate that this provides an earlier detection of a deteriorating situation than the traditional discrete time methods. Theoretical approximations to the average run length have also been derived for both discrete and continuous cases.

104 Longitudinal Data and Survival Data Modeling ● ✪

Biometrics Section, Section on Health Policy Statistics

Monday, July 30, 8:30 am–10:20 am

Joint Modeling of Longitudinal and Survival Data Using Penalized Corrected Likelihood

❖ Bo Hu, The Cleveland Clinic, 6811 mayfield Rd, Apt 785, Mayfield Heights, OH 44124, *hub@ccf.org*; Liang Li, The Cleveland Clinic; Tom Greene, The Cleveland Clinic

Key Words: Joint model, corrected likelihood

Joint modeling approach has been widely used for analyzing longitudinal data and survival data simultaneously. Many existing methods rely on certain algorithms to eliminate shared frailty parameters that link the longitudinal information and the survival information. We develop a penalized corrected likelihood method for analyzing such data. The corrected likelihood is a full likelihood, and is a concave function of the parameter, which ensures computational convergence.

Analysis of Longitudinally Measured Multivariate Biomarkers and Left-Censored Data

❖ Ghideon S. Ghebregiorgis, University of Pittsburgh, 2717 Cathedral of Learning, Pittsburgh, PA 15260, *ghg2@pitt.edu*; Lisa Weissfeld, University of Pittsburgh

Key Words: multivariate mixed model, Biomarkers, Longitudinal analysis, Left Censoring

Our objective is to analyze longitudinally measured multivariate biomarkers obtained from a large cohort study of patients who enrolled in a study examining markers for the development of sepsis in a cohort of subjects with community acquired pneumonia. For continuous outcomes, estimation is based on a parametric multivariate linear mixed model for the longitudinally measured biomarkers (Thiebaut et al. 2005).For bivariate left censored biomarkers, such as interleukin-6(IL-6)and interleukin-10(IL-10), an extension of this method based on the Tobit regression model will be used for the inference. The multivariate modeling approach is evaluated and compared with the univariate approaches through empirical findings and simulation studies. The simulation study will be used to measure the sensitivity of the performed analysis by considering various variance-covariances structures of the biomarkers

Modeling Longitudinal Biomarker Data with Multiple Assays That Have Different Known Detection Limits

❖ Paul Albert, National Cancer Institute, 6130 Executive Blvd Room 8136, Bethesda, MD 20906, *albertp@mail.nih.gov*

Assays to measure biomarkers are commonly subject to large amounts of measurement error and known detection limits. We propose a joint modeling approach for analyzing repeated measures of multiple assays when these assays are subject to measurement error and different known lower detection limits. A commonly used approach is to perform an initial assay with a larger lower detection limit on all repeated samples, followed by only performing a second more expensive assay when the initial assay value is below its lower detection limit. We show how simply replacing the initial assay measurement with the second assay measurement may be a biased approach and investigate the performance of the joint model in this situation. We evaluate different designs and illustrate the methodology with a study examining the use of a vaccine in treating macaques with simian immunodeficiency virus (SIV).

Regression Analysis of Multivariate Recurrent Event Data with Time-Varying Covariate Effects

❖ Liang Zhu, University of Missouri-Columbia, 3001 S Providence Road, 11G, Columbia, MO 65203, *lzkn7@mizzou.edu*

Key Words: event history study, longitudinal follow-up studies, marginal models, recurrent event data, time-varying coefficients

Recurrent event data occur in many fields and many approaches have been proposed for their analyses. However, most of the available methods only allow time-independent covariate effects and sometimes this may not be true as, for example, a treatment effect may decrease with time. In this paper, we consider the analysis of multivariate recurrent event data in which some covariate effects may be time-dependent. For the problem, some marginal models for the underlying counting processes are presented and an estimating equation-based inference procedure is presented. The asymptotic properties of the proposed estimates are established and their finite sample properties are evaluated through simulation studies. We also apply the approach to a set of real data.

Analysis of Longitudinally Measured, Left-Censored Biomarkers of Severe Sepsis with Dropout, and Death

❖ M. Abdus Sattar, University of Pittsburgh, 120 Ruskin Ave, Apt 122, Pittsburgh, PA 15213, *mas196@pitt.edu*; Lisa Weissfeld, University of Pittsburgh

Key Words: Longitudinal censored data, tobit regression, sensitivity analysis

Our goal was to address the problem of analyzing longitudinally measured left censored biomarker data when subjects are lost to data collection for multiple reasons. We used an un-weighted and weighted tobit regression method for interleukin-6 (IL-6) and interleukin-10 (IL-10) left censored biomarkers data; where the weight for the weighted method accounted for the various reasons of dropout. For IL-6, analysis of all of the predictors: race, steroid use, score of organ failure assessment (SOFA) and Pneumonia Severity Index (PSI), are significant though the standard errors of the estimated coefficients for the weighted tobit regression are smaller. For the IL-10 analysis, all of the above mentioned predictors are significant except PSI in the un-weighted tobit regression. The standard errors for the estimated regression coefficients are higher in the weighted tobit regression models.

Semiparametric Joint Modeling of Longitudinal and Time-to-Event Data Using P-Spline: A Penalized Likelihood Approach

❖ Wen Ye, University of Michigan, M4073 SPH II, 1420 Washington Heights, Ann Arbor, MI 48109, *wye@umich.edu*; Xihong Lin, Harvard School of Public Health; Jeremy Taylor, University of Michigan

Key Words: Joint Modeling, Survivial, Longitudinal, P-spline, Penalized Likelihood

Longitudinal studies in medical research often generate both repeated measurements of biomarkers and possibly censored survival data. Recently several joint models have been developed to deal with the challenges arising in this type of data. Commonly, in joint models, the longitudinal covariate is modeled by a linear mixed model. However, in some cases, the longitudinal covariate time trajectory is not linear. We propose a joint model using penalized cubic B-spline to accommodate the non-linear trajectory of longitudinal covariate measurements. To ease computation the estimation procedure is maximizing a penalized joint likelihood generated by a Laplace approximation of the joint likelihood, which combines the likelihood

of the longitudinal data and the partial likelihood of the time-to-event data. Properties of parameter estimators are investigated in simulation studies.

Joint Analysis of Multiple Longitudinal Outcomes and a Time-to-Event Using a Nonlinear Latent Class Approach

❖Cecile Proust-Lima, University of Michigan, Department of Biostatistics SPH, 1420 Washington Heights, Ann Arbor, MI 48109-2029, *cproustl@umich.edu*; Helene Jacqmin-Gadda, INSERM, U875; Jeremy Taylor, University of Michigan

Key Words: Mixed model, conditional independence assumption, joint analysis, maximum likelihood, latent class, longitudinal data

We propose a joint model for exploring the association between correlated longitudinal non-Gaussian markers and a time-to-event. A longitudinal latent class model describes the different latent classes of evolution of the latent process underlying the markers, and a proportional hazard model describes the risk of event according to the latent classes. The latent process is linked to the quantitative non-Gaussian markers by using nonlinear transformations which include parameters to be estimated. Depending on whether the function of risk is a parametric or semi-parametric function, a maximum likelihood estimation or a penalized likelihood approach is used. We propose a posterior residual analysis to evaluate the conditional independence hypothesis given the latent classes and apply the methodology in the context of cognitive aging.

105 Statistical Issues in Oncology Trials ●

Biopharmaceutical Section, Biometrics Section

Monday, July 30, 8:30 am–10:20 am

Modified Simon's Two-Stage Design with a Control Group

❖Sunil K. Dhar, New Jersey Institute of Technology, Department of Mathematical Sciences, 323 Martin Luther King Jr. Boulevard, Newark, NJ 07102, *dhar@njit.edu*; Farid Kianifard, Novartis Pharmaceuticals; Michael Chen, TCM Groups, Inc.

Key Words: Phase II design, Oncology

Simon's 2-stage design is a single-arm design, in which null hypothesis (H0) that response rate, p, is at most p0 is tested against alternative (H1) that p is at least p1. If there are fewer than k1 responders out of n1 enrolled in Stage 1, trial is terminated to accept H0. If at least k1 responders are observed, n2 additional subjects are enrolled in Stage 2. If there are fewer than k responders at the end of Stage 2, H0 is accepted; otherwise, H1 is accepted. For given Type I and II errors, design parameters n1, n2, k1, and k are estimated to meet certain optimality criteria. Without a control arm, results could be unreliable, especially when p0 and p1 values are unrealistic. We propose a 2-stage design with a control arm, and provide a SAS program to calculate optimal solutions for sample size and critical values at Stages 1 and 2 for given Type I and Type II errors, if they exist.

Censored Data Within Phase III Oncology Studies Using Independent Radiology Review (IRC) Assessments

❖Pralay Mukhopadhyay, Bristol-Myers Squibb Company, 5 Research Pkwy, PO Box 5100, Wallingford, CT 06460, *pralay.mukhopadhyay@bms.com*; Thomas Kelleher, Bristol-Myers Squibb Company

Key Words: IRC, censoring, discordance, power

Use of IRC determined assessments are common in oncology trials where the primary endpoint is progression free survival (PFS). In such studies disagreement between the study investigator and the IRC assessments of tumor progression may occur. When a patient comes off study due to an investigator based progression which the IRC disagrees with, progressions by IRC are censored. However contribution to PFS as a death event may occur which can bias the estimates of the medians and impact the overall results. Several methods of analysis to address this discordance will be explored through simulation studies.

Estimate of Treatment Effect with Adjustment for Post-Discontinuation Treatment Resumption and Crossover and Other Therapies in Oncology Clinical Trials

❖Shengyan Hong, Eli Lilly and Company, Indianapolis, IN 46285, *hongsh@lilly.com*; Yanping Wang, Eli Lilly and Company

Key Words: Treatment effect adjustment, post-discontinuation therapy, treatment resumption and crossover, proportional hazards model, oncology clinical trial

In oncology clinical trials of survival comparison between an experimental therapy and a best supportive care or an active control, it is clinical reality that many patients will receive additional therapies during post-discontinuation follow-up period. It is only natural to assume that the post-discontinuation therapies (PDTs) will have impact on patients' survival outcome and its effect should be adjusted for when estimating and comparing treatment effect. What makes this adjustment even more challenging is that some patients in one or both treatment arms could resume their assigned treatment or switch treatment as part of PDT. We propose a method to adjust for post-discontinuation treatment resumption and crossover as well as other therapies using Cox proportional hazards model approach. An example from a phase III clinical trial will be used to illustrate the proposed methodology.

A Two-Stage Multinomial Randomized Selection Design in Phase II Oncology Trials

❖Zhiping Sun, Merck & Co., Inc., 1258 Jarvis Ln, Lansdale, PA 19446, *linda_sun@merck.com*

Key Words: Randomized selection design, multinomial, two-stage

The objective of a phase II study in oncology is to evaluate whether an agent, at the doses determined by phase I trials, is worth of further investigation in phase III. Since several maximum tolerated doses corresponding to different dosing schedules may arise from phase I trials, a randomized selection design in phase II is more popular than a single arm design. Because progression-free rate is more related to the clinical endpoint(s) in phase III, a phase II design based on multiple outcomes is more reliable than a design based on the single conventional endpoint response rate. These concerns motivate the proposal and application of a two-stage randomized multinomial selection design in phase II oncology trials.

An Adaptive Two-Stage Design for Active-Controlled Phase II Oncology Trial

❖Xiaoming Li, sanofi-aventis, 860 Fieldcrest Ct, Blue Bell, PA 19422, *lixiaomi@gmail.com*

Key Words: oncology, two-stage design, adaptive, active-controlled

Typically phase II cancer trial is single-arm with tumor response rate as the primary endpoint, and two-stage designs are used. However, in many situations investigators are poorly served by such conventional methodology, primarily for two reasons: 1) there might be high uncertainty in targeting the alternative hypothesis, and 2) it is impossible to estimate the treatment effect as compared to historical control, since the treatment effect is confounded with differences in trials. An adaptive two-stage design with an active control is proposed to tackle the problems, in which the information from both treatment and control arms in the first stage is utilized to either reject or continue testing the study treatment with either an optimistic or a skeptic target response rate, based on outcomes from stage I. The characteristics of the proposed design are evaluated by simulation studies.

Optimal Three-Outcome Designs for Screening Trials of New Agents

❖Vandana Mukhi, New York University, 60 Sunflower Ridge Road, South Setauket, NY 11720, *vandana26@yahoo.com*; Yongzhao Shao, New York University

Key Words: Optimal Design, Three-outcome, Screening Trials

Recent interests in cancer vaccine development have prompted the need to test a series of agents to screen out promising ones for further study. A typical but major challenge is to use a limited number of patients and other resources to screen a large number of agents or treatments and identify promising ones for further study. Several optimal screening trial designs have been proposed in hypothesis testing framework with two possible outcomes, either reject the null hypothesis or do not reject the null hypothesis. Further extensions have also been proposed in the presence of distributional information on response rates. We further extend the optimal two-stage design to three-outcome framework, where the design allows for three possible outcomes: reject null hypothesis, do not reject the null hypothesis or inconclusive.

A Hybrid of Traditional Design and Isotonic Regression Method To Find MTD in Phase I Cancer Clinical Trials

❖Robert Bigelow, sanofi-aventis, 9 Great Valley Parkway, Malvern, PA 19355, *robert.bigelow@sanofi-aventis.com*; Chunpeng Fan, University of Wisconsin-Madison; Richard Wu, sanofi-aventis; Hui Quan, sanofi-aventis

Key Words: maximum tolerated dose, continual reassessment method, isotonic regression, phase I study designs

Current methods for estimating Maximum Tolerated Dose (MTD) include the traditional 3+3 design (TD), an accelerated traditional design (TDacc), continual reassessment methods (CRM), and isotonic regression (IR). Each method has its own limitation. The proposed 'hybrid' methods first use TDacc to quickly approach the neighborhood of MTD and then switch to IR. They use only the monotone dose response assumption for toxicity, as with IR, and do not assume an explicit dose-toxicity model. Results from simulation studies demonstrate that these hybrid methods can perform better than the CRM methods when the true dose-toxicity model is unknown. They also need fewer patients compared to IR and produce more accurate estimates compared to TD. A numerical example is used to illustrate the application of the methods.

106 Mixed Effect Models and Missing Data ●

ENAR, Biometrics Section

Monday, July 30, 8:30 am–10:20 am

A Pattern Mixture Model To Analyze Longitudinal Quality-of-Life Data with Nonignorable Dropout

❖Qinfang Xiang, Endo Pharmaceuticals, 100 Endo BLVD, Chadds Ford, PA 19317, *xiang.qinfang@endo.com*; Suna Barlas, Endo Pharmaceuticals

Key Words: Pattern mxiture model, Quality of Life, Missing not at random, Informative dropout, Longitudinal data analysis

The analysis of longitudinal data that suffer from missing not at random poses a statistical challenge. Inappropriate handling of the missing data that are of non-ignorable type can result in biased estimates and lead to incorrect conclusions. This study examines QoL data from a randomized clinical trial comparing the efficacy of Lidocaine Patch 5% with celecoxib 200mg in treating pain from Osteoarthritis of the knee. The QoL assessments are based on the pain interference on activities of daily living using Question9 of Brief Pain Inventory and Beck Depression Inventory. The study suffered from a significant amount of missing data (49%). A random effects pattern mixture model is applied to compare the two treatment groups in terms of the QoL indicators. This model is compared to those based on data that used the last observation carried forward, complete cases, and all observed cases.

Likelihood Ratio Hypothesis Testing in the Presence of Incomplete Data

❖Sergey Tarima, Medical College of Wisconsin, 8701 Watertown Plank Rd, H2460, Milwaukee, WI 53226, *starima@mcw.edu*; Albert Vexler, National Institutes of Health

Key Words: hypothesis testing, most powerful test, missing data

The main enemies of hypothesis testing procedures are (1) small sample sizes, (2) excessive nonresponse rate, and (3) high dimensionality. Their adverse effect can be reduced if additional information or guess about the distribution of alternatives can be incorporated into test statistics. If such information is too general (for example, only a parametrical family can be specified), this distribution is estimated from a principle of the most powerful testing. We investigate power properties of the proposed tests in the presence of the mentioned enemies and compare them with maximum likelihood estimators. Theoretical results, results of simulation studies and a missing data example are presented in the manuscript.

Principal Component Analysis for Multivariate Binary Data

❖Seokho Lee, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, *leesh@stat.tamu.edu*

Principal component analysis (PCA) is frequently used as a powerful tool for dimension reduction in multivariate data analysis. One of its disadvantages is that PCA is applicable only to continuous data. There is little literature on PCA-like dimension reduction methods for multivariate binary data which are commonly observed in various application fields. The nature of binary data circumvents a direct application of PCA to multivariate binary data. In this presentation, we propose an extension of PCA to handle multivariate binary variables by using Bernoulli distribution and low-rank approximations of matrices of latent variables. Illustrations of application to real and synthetic datasets are presented.

Overdispersion in Multilevel Correlated Data

❖Trent Lalonde, Arizona State University, Department of Mathematics and Statistics, Tempe, AZ 85287, *lalonde@mathpost.asu.edu*; Jeffrey R. Wilson, Arizona State University

Key Words: quasilikelihood, censoring, heterogeneity, extravariation, random effects

A common difficulty in regression problems with the exponential family response variables is overdispersion. While complex designs such as a multistage stratified cluster design greatly enhances the feasibility of data collection, it results in differential probabilities of selection as would not be the case if simple random sampling approach was employed. This phenomenon known as overdispersion has been examined through different approaches by several researchers. However, most articles either account for overdispersion caused by lack of independence among units or variation among responses within groups. This paper examines random effect mixed models that accounts simultaneously as well as at multilevels data.

Multiple Imputation Based on Functional Principal Components Analysis for Sparse Longitudinal Data

❖Szu-Ching Tseng, University of California, Davis, 325 East 8th St 7, Davis, CA 95616, *sctseng@ucdavis.edu*; Xiaowei Yang, University of California, Davis; Hao Zhang, University of California, Davis

Key Words: Sparse Longitudinal Data, Missing Data, Functional Principal Componenets Analysis, Linear Mixed-Effects Models

In biomedical research, due to unbalanced design with possibly large numbers of missing values, sparse longitudinal data are popularly seen in statistical practice. Recently, several solutions based on principal components analysis (PCA) for functional data have been developed. This paper aims to extend these methods within the framework of multiple imputation. By applying the PCA models to the sparse data with a chosen common time grid, complete data sets can then be generated. Each of these balanced data sets can be finally analyzed using standard longitudinal models such as random-effects models. By applying this imputation strategy to both simulated data sets and a hemodialysis vascular access data set, the performance of it is evaluated via comparing with the analysis using linear mixed-effects models to the original sparse data.

The Inclusion of Covariate Information in a Random Effects Agreement Model

❖Kerrie Nelson, University of South Carolina, Dept of Statistics, 1523 Greene St U of South Carolina, Columbia, SC 29208, *kerrie@stat.sc.edu*; Don Edwards, University of South Carolina

Key Words: agreement, generalized linear mixed model, kappa

When modeling agreement between two or more raters, it can be informative to identify factors which may influence the agreement process under study. For example, in the classification of a mammogram for the presence/absence of breast cancer, knowledge of the clinical history of the woman may impact the classification made by a rater, and in the agreement between a number of raters. In this talk, we describe how to include relevant covariate information in a random effects model for agreement, when binary classifications are of interest, and how a proposed model-based summary statistic can be used to assess the levels of agreement between raters in such a setting.

Maximum Likelihood Methods for Nonignorable Missing Data

❖Yan Zhou, University of Michigan, Dept of Biostatistics, Ann Arbor, MI 48105, *yzhouz@umich.edu*; Roderick J. Little, University of Michigan; John D. Kalbfleisch, University of Michigan

Key Words: nonignorable missing, maximum likelihood, patter-mixture model, factored likelihood

Missing values are common issues in empirical studies. The performance of the methods analyzing missing data strongly depends on the missing-data mechanism. If missing mechanism is ignorable, the likelihood based inferences will only depend on the observed data. For two variables X and Y where both of X and Y are missing, however, missing-data mechanism may not be ignorable, since the missingness of Y can possibly depend on the values of X which are missing. We propose a nonignorable missing-data mechanism, in which X is MCAR and Y is MAR given the value and missingness of X. The non-iterative maximum likelihood estimates exist and data are excluded for estimating certain parameters. Extensions of this type of mechanism will be also discussed.

107 Random Matrix Theory and High-Dimensional Inference ✪

IMS, Biometrics Section

Monday, July 30, 8:30 am–10:20 am

Effect of Mean on Variance Function Estimation in Nonparametric Regression

❖Lie Wang, University of Pennsylvania, 400 Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, *liewang@wharton.upenn.edu*; Lawrence D. Brown, University of Pennsylvania; Tony Cai, University of Pennsylvania; Michael Levine, Purdue University

Key Words: Minimax estimation, nonparametric regression, variance estimation

Variance function estimation in nonparametric regression is considered and the minimax rate of convergence is derived. We are particularly interested in the effect of the unknown mean function on the estimation of the variance function. Our results indicate that, contrary to the common practice, it is often not desirable to base the estimator of the variance function on the residuals from an optimal estimator of the mean. Instead it is desirable to use estimators of the mean with minimal bias.

The Largest Root Test in Multivariate Analysis: Nonasymptotic Accuracy of the Tracy-Widom Approximation

❖Iain Johnstone, Stanford University, School of Humanities and Sciences, Building One Main Quad, Stanford, CA 94305, *imj@stanford.edu*

Key Words: multivariate analysis, canonical correlations, MANOVA, mutiple responses, Tracy-Widom, random matrices

The largest root test is a staple of classical multivariate analysis, in canonical correlation analysis, MANOVA, multiple response linear regression and so on. It is based on the null distribution of the largest eigenvalue of A(A+B)^{-1}, where A and B are independent Wishart matrices on p variables. With appropriate centering and scaling, the Tracy-Widom law provides an approximation to the null distribution. The approximation comes from asymptotics in which the number of variables, p, is large. The purpose of this talk is to show—by numerical comparison with exact results—that the approximation may be quite adequate for applied work at conventional significance levels for small p, even as low as 2.

Statistical Eigen-Inference from Large Wishart Matrices

❖Raj Rao, Massachusetts Institute of Technology, 195 Binney Street Apt 2227, Cambridge, MA 02142, *raj@mit.edu*

Key Words: random matrices, eigen-inference, Wishart matrices

The asymptotic behavior of the eigenvalues of a sample covariance matrix is described when the observations are from a zero mean multivariate (real or complex) normal distribution whose covariance matrix has population eigenvalues of arbitrary multiplicity. In particular, the asymptotic normality of the fluctuation in the trace of powers of the sample covariance matrix from the limiting quantities is shown. Concrete algorithms for analytically computing the limiting quantities and the covariance of the fluctuations are presented. Tests of hypotheses for the population eigenvalues are developed and a technique for inferring the population eigenvalues is proposed that exploits this asymptotic normality of the trace of powers of the sample covariance matrix. Numerical simulations demonstrate the robustness of the proposed techniques in techniques in high-dimensional settings.

New Optimality Test for Branch-and-Bound–Based Comprehensive Subset Search

❖ Xuelei Ni, Kennesaw State University, 1000 Chastain Rd, Mail Drop 1204, Kennesaw, GA 30144, *xni2@kennesaw.edu*; Xiaoming Huo, Georgia Institute of Technology

Key Words: subset selection, branch-and-bound, leaps and bounds, optimality test, inverse tree, pair tree

In subset selection in regressions, branch-and-bound algorithms and the Furnival and Wilson (1974) test have been combined to search all subsets. Furnival and Wilson test reduces the number of subsets that needed to be computed, and is completely based on a monotonic property of the residual sum of squares. We design a new test, which brings in other quantities (e.g., the least square coefficients and the residuals) into consideration. The new test supplements the Furnival-Wilson test. Simulations demonstrate improvements brought by the newly proposed test. We describe a branch-and-bound structure|pair tree|within which the new optimality test can be easily integrated. The pair tree is comparable with other contemporary branch-and-bound designs, for which pointers are given.

High-Dimensional Classification Using Features Annealed Independence Rule

❖ Yingying Fan, Princeton University, Department of ORFE, Princeton, NJ 08544, *yingying@princeton.edu*; Jianqing Fan, Princeton University

Key Words: Classification, feature extraction, high dimensionality, independence rule, misclassification rates

High-dimensional classification arises frequently in contemporary statistical problems. The impact of dimensionality on classifications is poorly understood. We first demonstrate that even for the independence classification rule, classification using all features can be as bad as random guessing due to noise accumulation in estimating population means in high-dimensional setting. In fact, we prove that almost all linear discriminants can perform as bad as random guessing. Thus, it is important to select a subset of important features, resulting in Feature Annealed Independence Rules. The conditions under which all important features can be selected by the two-sample t-test are established. The choice of the optimal number of features is proposed based on an upper bound of the classification error. Simulation studies and real data analysis support our theoretical results.

On the Minimax Risks for Estimation of a Location Parameter from Sums

❖ Mokshay Madiman, Yale University, 24 Hillhouse Avenue, New Haven, CT 06511, *mokshay.madiman@yale.edu*; Andrew Barron, Yale University; Abram Kagan, University of Maryland; Tinghui Yu, University of Maryland

Key Words: Minimax risk, Pitman estimator, Distributed estimation, Fisher information, Information inequalities

Suppose there are N random variables of interest, and each user has access to observations from the location family generated by a particular subset sum. For a given class of subsets, the variances of the Pitman estimates obtained by the corresponding class of users is related to the variance of the Pitman estimate obtained by a user who only sees observations from the location family generated by the total sum. This relationship is made explicit through a new inequality, which strengthens various inequalities for the Fisher information of a sum due to Stam (1959), Artstein, Ball, Barthe and Naor (2004), and Madiman and Barron (2006). In addition to providing insights on distributed estimation, the new result relates the minimax risks (in terms of mean squared error) incurred for estimating a location parameter by users who have access to different sets of observations.

108 Nonparametric Statistics and Mixture Models ✪

IMS, Section on Nonparametric Statistics

Monday, July 30, 8:30 am–10:20 am

Set and Set Properties Estimation

❖ Bruno Pelletier, University Montpellier II, Place Eugene Bataillon, Montpellier, 34095 France, *pelletier@math.univ-montp2.fr*; Gerard Biau, University Montpellier II; Benoit Cadre, University Montpellier II

Key Words: nonparametric statistics, support estimation, clustering, set estimation

We study the problems of reconstructing a set S and of estimating one its topological characteristic, namely its number of connected components, from random points of S drawn from some probability measure. We focus on the certainly most simple set estimator defined as the union of balls centered at the random points. Additionally, we propose a graph-based estimator of the number of connected components of S. Using tools from Riemannian geometry, and under mild analytic conditions on the underlying density of the data, we derive the exact rate of convergence of this set estimator and prove the consistency of the estimator of the number of connected components. Statistical applications include density support estimation and estimation of the number of clusters in data partitioning.

Distribution Theory of Order Statistics of Concomitants Subsets and Applications

❖ Ke Wang, The Ohio State University, 5961 Dartshire Blvd, Dublin, OH 43016, *kewang@stat.ohio-state.edu*; Haikady Nagaraja, The Ohio State University

Key Words: order statistics, concomitants of order statistics, conditional distribution, limiting distribution

Let (Xi, Yi), i = 1, ..., n be a random sample of size n from an absolutely continuous bivariate distribution F(x, y). We study the finite-sample and asymptotic joint distribution of (V(s:m),W(t:n-m)), where V(s:m) is the s-th order statistic of the concomitants subset {Y[i:n], i=n-m+1,...,n}, and W(t: n-m) is the t-th order statistic of the concomitants subset {Y[j:n], j =1, ..., n-m}. The results are applied to study the probability that at least s of the Y concomitant values, {Y[n-m+1:n], Y[n-m+2:n], ..., Y[n:n]}, are among the top k values of the entire Y -sample. Such events are of interest in selection procedures as described in Yeo and David (1984). We also apply the results to study the power of the two-stage designs for gene-disease association studies as discussed in Satagopan et al. (2002) and Satagopan et al. (2004).

Differentiated Logdensity Approximants

❖ Serge Provost, The University of Western Ontario, Dept of Stat and Act Sciences WSC 262, London, ON N6A5B7 Canada, *provost@stats.uwo.ca*

Key Words: Density approximation, Percentiles, Moment-based methodologies, Orthogonal polynomials, Computational Statistics

An alternative density approximation methodology according to which the first derivative of the logarithm of the density approximant is expressed as a ratio of polynomials, is being proposed. The polynomial coefficients of the numerator and denominator of the rational function are determined by solving the system of linear equations which results from matching the moments of the approximant to those of the distribution being approximated. As will be explained, such approximants possess numerous properties of interest. The results will be compared in specific applications to those obtained from the saddlepoint technique and methodologies based on orthogonal polynomials and Edgeworth expansions. Several illustrative examples will be provided.

On a Grouping Method for Constructing Mixed Orthogonal Arrays

❖ Chung-yi Suen, Cleveland State University, Department of Mathematics, Cleveland State University, Cleveland, OH 44115-2214, *c.suen@csuohio.edu*

Key Words: finite field, finite projective geometry, mixed orthogonal array, flat, tight

Mixed orthogonal arrays of strength two and size s^{mn} are constructed by grouping points in the finite projective geometry PG(mn-1,s). PG(mn-1,s) can be partitioned into (s^{mn}-1)/(s^n-1) (n-1)-flats such that each (n-1)-flat is associated with a point in PG(m-1,s^n). A set of (s^t-1)/(s-1) points in PG(m-1,s^n) is called a (t-1)-flat over GF(s) if it is isomorphic to PG(t-1,s). If there exists a (t-1)-flat over GF(s) in PG(m-1, s^n), then we can replace (s^t-1)/(s-1) s^n-level columns in an orthogonal array by (s^n-1)/(s-1) s^t-level columns and obtain a mixed orthogonal array. Many new mixed orthogonal arrays can be obtained by this procedure. New orthogonal arrays of sizes 256, 512, and 1024 are also obtained using PG(6,2), PG(7,2), and PG(8,2) respectively.

Labeling Issue in Finite Mixture Model: A Frequentist View

❖ Daeyoung Kim, The Pennsylvania State University, 326 Thomas Building, University Park, PA 16802, *dzk123@psu.edu*; Bruce G. Lindsay, The Pennsylvania State University

Key Words: Finite mixture model, Labelling non-identifiability, Asymptotic identifiability, Parametric bootstrap

Our research addresses the problems caused by labeling nonidentifiability in a parametric finite mixture model from the frequentist view. We propose a nonparametric method to see if the labels on the parameters are well defined so as to be useful in deciding Monte Carlo (bootstrap) standard errors for the parameters. Although there is a form of asymptotic identifiability for consistent labeling on the parameters, this will not work well when the components densities are not well separated or the sample size is not large. The nonparametric approach we investigate provides information on the labeling identifiability for parameters by constructing a data set that contains each bootstrapped parameter estimate multiple times, one for each possible labeling. If this data set breaks into well-separated clusters, then we consider asymptotic identifiability to be reasonable.

Using Copula To Study Dependence in Mixed Distributions

❖ Magdalena Niewiadomska-Bugaj, Western Michigan University, Everett Tower, Kalamazoo, MI 49008, *m.bugaj@wmich.edu*

Key Words: zero inflated data, mixed distributions, dependence, copula

Nonnegative continuous distributions with a positive probability at zero are often used for modeling in the medical, environmental and other research areas. Most of the existing statistical techniques cannot be used for analysis of such data because the necessary assumptions (e.g., continuity, normality) are violated. This paper is devoted to studying dependence in nonnegative variables with a positive probability at zero. Through the grade transformation the mixed distribution becomes a copula and the dependence can then be studied accordingly. We will focus on implications, advantages and limitations of the proposed approach.

109 Bayesian Case Studies ●✪

Section on Bayesian Statistical Science

Monday, July 30, 8:30 am–10:20 am

Bayesian Real-Time Model for Ovulation Prediction

❖ Lemuel Moye, University of Texas School of Public Health, 3607 Latma Drive, Houston, TX 77025, *moyelaptop@msn.com*

Key Words: Bayesian, ovulation, physicians

Women of child bearing age commonly require accurate knowledge of their ovulation time. Working with fertility specialists, a Bayesian model for the real time prediction of the day of ovulation was developed. This model, based on the woman's prior ovulation history, her minimum and maximum basal body temperature (BBT) during the cycle, and her current BBT, provides the probability distribution that today is the ovulation day. This real time assessment is preferable to other models that, while providing ovulation assessments, must have the complete cycle to make their retrospective assessment. Designed to be used by both physicians and patients, this computing tool is easily programmed onto a PDA or website. Examples are provided.

Bayesian Multiple Outcomes Models: Benefits and Challenges

❖ Sally W. Thurston, University of Rochester, Department of Biostatistics and Computational Bio, 601 Elmwood Ave. Box 630, Rochester, NY 14642, *thurston@bst.rochester.edu*; David Ruppert, Cornell University

Key Words: mercury, mixed model, multiple outcomes, random effects

Understanding the relationship between prenatal mercury exposure and neurodevelopment in children is of great interest to many practitioners. Typically, analyses rely on separate models fit to each outcome. If the effect of exposure is very similar across outcomes, separate models lack power to detect a common exposure effect. Furthermore, the outcomes cluster into broad domains and domain-specific effects are also of interest. We fit a Bayesian model which allows the mercury effect to vary across outcomes, while allowing for shrinkage of these effects within domains, and to a lesser extent between domains. We will discuss the benefits and challenges of fitting this model within a Bayesian framework, and present results of the model applied to multiple outcomes measured in children at 9 years of age in the Seychelles.

Analyzing Pressure Ulcer Development of 36 Nursing Homes Using Bayesian Hierarchical Modeling

❖ Jing Zhang, University of Missouri-Columbia, 510 High Street Apt 316, Columbia, MO 65201, *jztw7@mizzou.edu*; Zhuoqiong (Chong) He, University of Missouri-Columbia

Key Words: Hierarchical, logistic regression, MCMC

Pressure ulcer development is an important measurement in judging the quality of nursing service, therefore it is an interesting research topic to find out the effect of nursing homes in the measurements of ulcer development rates. A lot of different methods have been used, including traditional logistic regression, Bayesian Hierarchical modeling and semiparametric approaches. In this report we use Bayesian Hierarchical Models and apply MCMC algorithm to predict the performance of 36 nursing homes in term of pressure ulcer development of the residents in these institutes. Bayesian approach provides posterior distributions besides estimates and predictions, therefore we could develop different criteria for detecting the nursing homes that are providing problematic nursing services. We also performed cross validation to evaluate the predictive ability of the Bayesian Hierarchic model.

Application of Bayesian Analysis in Review of Adverse Event for Product Quality Issues

❖ Harry Yang, MedImmune, Inc., One MedImmune Way, Gaithersburg, MD 20878, *yangh@medimmune.com*; Lanju Zhang, MedImmune, Inc.; Iksung Cho, MedImmune, Inc.

Key Words: Bayesian analysis, Adverse event, Prior, Monte Carlo simulation

A pharmacovigilance system capable of identifying product lots with quality issues through analyzing post-marketing safety data is very important. Although some researchers have proposed safety data review procedures to signal potentially problematic lots, few of the proposals were based on sound statistical methodologies. We present a Bayesian and frequentist approaches to detecting out-of-trend AEs. The number of AEs is described through a mixture model, taking on binomial and Poisson distributions, with a probability of p and 1 - p, respectively. Assuming parameters p and l (Poisson mean) have Beta and Gamma priors, respectively, the probability for the number of AEs of the current lot under evaluation to exceed its observed value is estimated. This estimate allows us to determine if the product lot associated with the AEs is of quality concerns. A method based on MLE is proposed.

Bayesian Analysis of Surveillance Data and Meteorological Data for Detection of Aerosol Releases of B. Anthracis Spores

❖ William Hogan, University of Pittsburgh, Suite M-183, Parkvale Bldg, 200 Meyran Avenue, Pittsburgh, PA 15260, *wrh@cbmi.pitt.edu*; Garrick Wallstrom, University of Pittsburgh; Gregory F. Cooper, University of Pittsburgh; Michael Wagner, University of Pittsburgh; Jean-Marc Depinay, University of Pittsburgh

Key Words: Biosurveillance, Statistical outbreak detection, Atmospheric dispersion, Anthrax, Emergency department data, Bayesian data analysis

Statistical analysis of data is a key component of biosurveillance. The newer types of data it monitors, such as emergency-department (ED) visits, have variable baseline activity and extracting the signal of an outbreak from this noise is difficult. We present the Bayesian Aerosol Release Detector (BARD), an algorithm to detect large, outdoor aerosol releases of B. anthracis spores as early as possible. BARD estimates the posterior probability of a release–and a posterior distribution over the release location, quantity, and time–from ED visit and meteorological data. It comprises a Bayesian model of inhalational anthrax and background disease and a deterministic atmospheric dispersion equation. In an evaluation, BARD detected simulated releases in a timely, sensitive, and specific manner. Its performance was much less sensitive to varying model parameters than to outbreak size.

Hierarchical Bayesian Markov Switching Models with Application to Predicting Spawning Success of Shovelnose Sturgeon

❖ Ginger Davis, University of Virginia, 151 Engineers Way, PO Box 400747, Charlottesville, VA 22904-4747, *gingerdavis@virginia.edu*; Scott Holan, University of Missouri-Columbia; Mark Wildhaber, U.S. Geological Survey

Key Words: hierarchical, Markov-switching, longitudinal, GARCH, Bayesian, eigenvalue

Spawning in sturgeon is linked to environmental patterns, rhythms, and cues. Due to the endangerment of sturgeon, efforts need to be made to increase recruitment. Little information is available on the biology and ecology of sturgeon to guide these efforts. It is not known where, when and under what conditions these species spawn in the Missouri River, and to what degree spawning is successful. Using measurements of biological variables associated with readiness to spawn as well as longitudinal behavioral data collected using telemetry and data storage device sensors, we introduce a hierarchical Bayesian model for predicting spawning success. This model uses an eigenvalue predictor from the transition probability matrix in a two-state Markov switching model with GARCH dynamics as a generated regressor in a linear regression model.

110 Nonparametric Testing and Model Validation

Section on Nonparametric Statistics

Monday, July 30, 8:30 am–10:20 am

A Generalized Test Statistic for Multinomial Goodness-of-Fit

❖ Sunil Mathur, University of Mississippi, Hume 325, University, MS 38677, *skmathur@olemiss.edu*

Key Words: Generalized, Goodness-of-fit, Asymptotic distribution

Several goodness-of-fit tests are available in literature for testing the goodness of fit of discrete multivariate data. We propose a unified analysis of goodness of fit using a family of statistics. Depending on the choice of parameters, the generalized test statistics results into some of the known statistics such as Chi-square, and likelihood ratio generalized test. The proposed test statistic is highly robust against extreme values and does not assume the distribution of parent population. The asymptotic distribution of the proposed test statistic, and the p-value function are discussed. The application of proposed method is shown by using a real-life data.

Model-Checking in Errors-in-Variables Regression

❖ Weixing Song, Kansas State University, Department of Statistics, Manhattan, KS 66502, *weixing@ksu.edu*

Key Words: Errors-in-Variables Model, Deconvolution Kernel, Minimum Distance, Lack-of-Fit

This paper discusses a class of minimum distance tests for fitting a parametric regression model to a class of regression function in the errors-in-variables model. These tests are based on certain minimized distances between a nonparametric regression function estimator and a deconvolution kernel estimator of the conditional expectation of the parametric model being fitted. The paper establishes the asymptotic normality of the proposed test statistics under the null hypothesis and that of the corresponding minimum distance estimators. The significant contribution made in this research is the removal of the common assumption in the literature that the density function of the design variable is known. A simulation study shows that the testing procedures are quite satisfactory in the preservation of the finite sample level and in terms of a power comparison.

A Lack-of-Fit Test for Partially Linear Models

❖ Chin-Shang Li, St. Jude Children's Research Hospital, Department of Biostatistics, 332 N Lauderdale St, Memphis, TN 38105, *chinshang.li@stjude.org*

Key Words: B-splines, Wald-type test

The nonparametric component of a partially linear model is parameterized by a linear combination of fixed knot cubic B-splines. The estimators of the model parameters are estimated by minimizing the penalized least-squares criterion. A Wald-type test based on the proposed estimation method is proposed for assessing the linearity of the nonparametric component. The

finite-sample property of the proposed test is illustrated, and the power performance of the test is studied through simulations.

Robust and Powerful Two-Sample Semiparametric Tests

❖Zhong Guan, Indiana University South Bend, Dept of Mathematical Sciences, South Bend, IN 46634, *zguan@iusb.edu*; Hongyu Zhao, Yale University

Key Words: Empirical likelihood, Two-sample test, Semiparametric test, Behrens-Fisher problem, Bootstrap, Power

A semiparametric model is constructed for testing the equality of parameters of two populations, such as means or medians. The semiparametric likelihood ratio based tests are studied and compared with some parametric and nonparametric ones. The numerical comparison shows that the semiparametric approach is more robust than the parametric tests and more powerful than the nonparametric tests in the sense that it has similar powers to some optimum parametric tests for some distributions, it is more powerful than two-sample t and Wilcoxon tests when the parameters are not location parameters. The validity of the bootstrap procedure for estimating p-values is also proved.

A Modified Mood's Rank Test for Comparing Two Crossing Distributions

❖Michael P. Anderson, Kansas State University, 2103 Prairie Field, Manhattan, KS 66502, *mpa@ksu.edu*

Key Words: quantile-preserving spread, two-sample test, two crossing distributions, linear rank test statistic

This paper proposes a modification to the traditional nonparametric Mood's test to carry out hypothesis tests for "quantile-preserving spreads" defined in Mendleson (1988). The test makes use of a procedure for estimating the crossing point of two distributions developed in a master's report by Zong (2004). Critical values for the proposed test are tabulated using Monte Carlo simulation. A simulation study is carried out to compare the proposed test with the classical rank tests developed by Mood, Ansari-Bradley, Seigel-Tukey, and Mendelson as well as a test proposed by Zong (2004). The use of the test is demonstrated using guinea pig data given in Doksum (1974).

Goodness-of-Fit Tests for Multinomial Logistic Regression

❖David Hosmer, University of Massachusetts Amherst, 128 Worcester Road, Stowe, VT 05672, *hosmer@schoolph.umass.edu*; Morten W. Fagerland, Ullevål University Hospital; Anna M. Bofin, Norwegian University of Science and Technology

Key Words: regression models, generalized linear models, fit tests, simulations

We examine goodness-of-fit tests for multinomial logistic regression. One is based on a g by outcomes contingency table. The test statistic (Cg) is obtained by calculating the usual Pearson chi-square statistic from this table. Simulations compare the properties of Cg to the Pearson chi-square test (X2) and its normalized test (z). The null distribution of Cg is approximated by the chi-square distribution with (g -2) ◊ (c-1) degrees of freedom. X2 is compared to a chi-square with n ◊ (c - 1) degrees of freedom, but shows erratic behavior. z adheres reasonably well to the standard normal distribution. Power simulations show that Cg and z have low power for a sample size of 100 observations, but satisfactory power for a sample size of 400. The tests are illustrated using data from a study of cytological criteria for the diagnosis of breast tumors.

Model Credibility Indices and Information Criteria

❖Jiawei Liu, Georgia State University, Department of Math and Scieces, Atlanta, GA 30303, *matjxl@langate.gsu.edu*; Bruce G. Lindsay, The Pennsylvania State University

Key Words: model selection, AIC, BIC, bootstrap, normality

A standard goal of model evaluation and selection is to find a model that approximates the truth well while at the same time is as parsimonious as possible. In this paper we emphasize the point of view that the models under consideration are almost always false, if viewed realistically. We investigate this issue in large samples by looking at two types of model credibility indices, which are designed to serve as one-number summary measures of model adequacy. We define these indices to be the maximum sample size at which samples from the model and those from the true data generating mechanism are nearly indistinguishable. We show that these definitions lead us to some new ways of viewing models as flawed but useful. We will again construct sample-size-index measures that assess the quality of models at different sample sizes, but in a model comparison setting.

111 Symbolic, Time Series, and Image Analysis I

Section on Statistical Computing, Section on Statistical Graphics, IMS

Monday, July 30, 8:30 am–10:20 am

Exact Properties of a New Test and Other Tests

❖Jie Peng, University of Louisiana at Lafayette, 326 A Cleveland Street, Lafayette, LA 70501, *jxp5159@ull.edu*; Kalimuthu Krishnamoorthy, University of Louisiana at Lafayette

Key Words: Conditional exact test, Mid-p value, Powers, Parametric bootstrap, size

The problem of testing equality of several binomial proportions is considered. An AU-test is proposed by extending a result for the two-sample case. Exact binomial distributions are used to evaluate Type I error rates of the usual chi-square test, an exact conditional test, a conditional test based on mid p-values and the AU-test are evaluated numerically. The AU test and the conditional test based on mid p-values control the Type I error rates well even for small samples whereas the exact conditional test is too conservative. The powers of the chi-square test, conditional test based on mid-p values, and the AU test are evaluated and compared. Power comparison shows that all three tests exhibit similar power properties when their sizes are within the nominal level. The AU-test practically behaves like an exact test even for small samples, and can be safely used for applications.

On a Moment-Based Test for Normality

❖Yihao Deng, Indiana University Purdue University Fort Wayne, Dept of Mathematical Sciences, 2101 E. Coliseum Blvd., Fort Wayne, IN 46835, *dengy@ipfw.edu*; Chand Chauhan, Indiana University Purdue University Fort Wayne

Key Words: Test, Normality, Moment

Normality is one of the important assumptions in many statistical analysis. Tests for normality including skewness test, kurtosis test have been extensively studied from different aspects. In this talk, we propose an alternative test based on several moments. Since moment generating function uniquely determines a distribution, the developed test has some desired features such as simple computation, ease of understanding and reasonable power. Some simulation and numerical results are also presented.

Testing of Hypothesis of a Structured Mean Vector for Multilevel Multivariate Data with Structured Correlations on Repeated Measurements

✻*Key Words:* autoregressive correlation structure, compound symmetry correlation structure, equicorrelated (partitioned) variance covariance matrix, maximum likelihood estimates, multi-level multivariate data, multivariate equicorrelation

We study the testing of hypothesis of a structured mean vector against an unstructured mean vector for multilevel multivariate data, where more than one response variable is measured on each experimental unit on more than one site at several time points under the assumption of multivariate normality. We provide the maximum likelihood estimates of the unknown population parameters for both autoregressive of order one and compound symmetry correlation structures on the repeated measures over time. We also provide the computation algorithms for calculating the test statistics. The proposed test is illustrated with four real datasets.

Resampling-Based Multiple Testing Procedure

❖Nasrine Bendjilali, Lehigh University, 23626 Magic Mountain PKWY, Apt 516, Valencia, CA 91355, *nab4@lehigh.edu*; Wei-Min Huang, Lehigh University

Key Words: false discovery rate, p-value, multiple hypotheses testing

This article introduces a method for multiple hypotheses testing which is based on the jump size between the potential "smallest null" p-value and the potential "largest alternative" p-value with the structure of resampling methods. The size of the jump is used to determine the stopping rule of the procedure and the resampling incorporates the covariance structure of the data without the need for any distributional assumptions. The proposed procedure seems to have fairly good control of FDR (the false discovery rate). Preliminary study also shows promising gain in power.

On an Efficient Algorithm for Boundary Detection

❖Tsung-Lin Cheng, National Changhua University of Education, Department of Math, Changhua, 500 Taiwan, *tlcheng@cc.ncue.edu.tw*

Key Words: change curve, random field

We propose a novel and efficient algorithm to detect the change-curve which separates a random field into two parts with distinct distributions. This algorithm, compared with the traditional cumulative-sums (CUSUM) estimations in the one-dimensional case (cf. Cheng 2006), not only simplifies the proofs of the consistency of the estimators of the change-point but also dramatically reduces the computation time without sacrificing too much precision. We extend this method to the random fields under mild conditions. To illustrate the algorithm, we perform some Monte Carlo simulations for different models and different changes.

A Geometric Feasible Direction Algorithm for Large-Scale Optimization with l1 Norm Constraint

❖Jian Zhang, Purdue University, 250 N. University St., West Lafayette, IN 47907, *jianzhan@stat.purdue.edu*

Key Words: lasso, constrained optimization, feasible direction, variable selection

Recently, a lot of interests in machine learning and statistics have been given to learning methods which can lead to sparse solutions. Among those learning methods, many can be formulated as an optimization problem with l1 norm constraint. We propose a feasible direction algorithm which utilizes the geometric structure of the feasible region and is computationally efficient. We present sufficient and necessary conditions for optimality, as well as its convergence properties. Numerical experiments are conducted to show that, compared to several other methods, it is very efficient, robust and accurate.

Impact of Censoring on Inference for the Regression Coefficient in a Bivariate Normal Model

❖Richard Linder, Ohio Wesleyan University, Delaware, OH 43015, *rslinder@owu.edu*

Key Words: Censoring, Correlation, Order Statistics, Concomitants of Order Statistics, Bivariate Normal Model, Regression

Suppose a random sample of bivariate normal data is subjected to Type II censoring on one of the variates, so that only a set of order statistics and their concomitants are observed. Inference about the correlation parameter and the regression coefficient is described here in this context. First, properties of the maximum likelihood estimator of the regression coefficient are described and compared with those of a range-based estimator from the literature. Two simple methods for approximating power of the test for independence in this model are proposed. Finally, based on these, a method is illustrated for determining sample sizes necessary to test for independence at specified power.

112 Spatial Methodology

Section on Statistics and the Environment

Monday, July 30, 8:30 am–10:20 am

On the Number of Clusters in a Dataset

❖Ahmed Albatineh, Nova Southeastern University, Division of MST, 3301 College Avenue, Fort Lauderdale, FL 33314, *albatine@nsu.nova.edu*

Key Words: Similarity index, Clustering algorithm, Circular data, Bivariate normal mixture, von Mises distribution, Correction for chance agreement

This paper develops a method for determining the number of clusters in a data set by studying the behavior of the corrected similarity indices between two clustering methods. The proposed method is applied to a simulated bivariate normal data and further extended for use in circular data. Its performance is compared to some criteria in SAS software. Overall, the proposed method was found to perform very well. The effect of cluster size, number of clusters, dimensionality, and covariance structure have been investigated. The method is not based on any distributional or data assumption which makes it widely applicable. The results using real data example is given.

Assessing Statistical Spatial Performance: An Application in Archaeology

❖Yuemei Wang, Emory University, Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322, *ywang27@sph.emory.edu*; Lance Waller, Emory University; Zev Ross, ZevRoss Spatial Analysis

Key Words: statistical spatial performance, Generalized Additive Models, receiver operating characteristic (ROC) curve

We investigate the problem of cluster detection when adjusting for covariates using Generalized Additive Models (GAM). GAMs are extensions of generalized linear models replacing some or all parametric terms with smooth functions. In addition, GAMs provide spatial smoothing for 2-dimensional data by using thin-plate spline regression. We apply GAMs to data from the Black Mesa archaeological project to identify clusters of early versus late Anasazi settlement sites while adjusting for exposure to rivers around those sites. We also evaluate the statistical power of the approach as a function of space. We generate the receiver operating characteristic

(ROC) curve associated with classifications of locations based on Monte Carlo simulations and compare the GAM results with results based on kernel density estimation of the early-to-late relative risk surface and SaTscan cluster statistics.

Mining Semantic Co-Location Patterns with Clustering Techniques

❖Bin Zhang, IBM China Research Laboratory, Building 19 Zhongguancun Software Park, Beijing, 100094 China, *zbin@cn.ibm.com*; Wen Jun Yin, IBM China Research Laboratory; Jin Dong, IBM China Research Laboratory; Ming Xie, IBM China Research Laboratory

Key Words: co-location mining, semantic co-location mining, spatial mining

Spatial co-location mining technologies are usually leveraged in broad applications to find subsets of spatial features frequently located together in spatial proximity, e.g. the frontage roads and highways in metropolitan road maps, and co-located services frequently requested together from mobile devices in location-based services. Traditional co-location mining approaches mainly focus on spatial attributes but consider less the internal attributes of objects. In this paper, we define a semantic-based interestingness measure to take into account internal attributes and extend the pure spatial co-location pattern to semantic co-location pattern. A clustering-based co-location pattern mining approach is also developed to discover those semantic co-location patterns. In the experiment, a real-world case study shows the proposed approach can effectively find semantic co-locations.

Spatial Modeling for Large Multivariate Environmental Data: Advancing Methods and Applications

❖Sudipto Banerjee, The University of Minnesota, 420 Delaware Street SE, A460 Mayo Bldg MMC303, Minneapolis, MN 55455, *sudiptob@biostat.umn.edu*; Andrew Finley, The University of Minnesota

Key Words: Bayesian modelling, Markov Chain Monte Carlo, Multivariate Processes, Spatial Predictions

Hierarchical spatial process models implemented through Markov Chain Monte Carlo (MCMC) for understanding scientific relationships, though flexible and versatile, involve expensive matrix decompositions rendering them infeasible for large spatial datasets. This computational burden is exacerbated in multivariate settings with several spatially dependent response variables. Here we propose to use a predictive process to model multivariate spatial data that projects process realizations to a lower-dimensional subspace thereby reducing the computational burden. The resulting predictive process models enjoy attractive theoretical properties along with greater modeling flexibility. We show how the predictive process adapts to multivariate nonstationary processes, with richer association structures. A computationally feasible template that encompasses these diverse settings will be presented.

A Multivariate Semiparametric Bayesian Spatial Modeling Framework for Hurricane Surface Wind Fields

❖Brian Reich, North Carolina State University, 2501 Founders Drive, 209H Patterson Hall, Raleigh, NC 27695-8203, *reich@stat.ncsu.edu*; Montserrat Fuentes, North Carolina State University

Key Words: Bayesian, nonparametric, multivariate, spatial, wind field, hurricane

Numerical ocean models are essential for creating storm surge forecasts for coastal areas. These models are driven primarily by the surface wind forcings. A new nonparametric multivariate spatial modeling framework is introduced combining data with physical knowledge about the wind fields to improve the estimation of the wind vectors. Our model builds on the stick-breaking prior, which is frequently used in Bayesian modeling to capture uncertainty in the parametric form of an outcome. The stick-breaking prior is extended to the spatial setting by assigning each location a different, unknown distribution, and smoothing the distributions in space with a series of kernel functions. This semiparametric spatial model is shown to improve prediction compared to usual Bayesian kriging methods for the wind field of Hurricane Ivan.

Maximum Likelihood for Spatially Correlated Discrete Data

❖Lisa Madsen, Oregon State University, Department of Statistics, 44 Kidder Hall, Corvallis, OR 97331, *madsenl@onid.orst.edu*

Key Words: spatial statistics, discrete data, copula

Techniques for estimating relationships between a spatially indexed response and a set of covariates are well-known when the response can be assumed Gaussian. For a discrete response, some authors have adapted Liang and Zeger's (1986) generalized estimating equations approach. These models typically employ a spatially correlated latent variable. Latent variable models place artificially low limits on correlations, which can lead to underestimating standard errors. I propose a copula-based maximum likelihood approach for spatially correlated discrete data. This approach allows modeling of correlations up to the theoretical limit, and, given appropriate regularity conditions, enjoys all the nice properties of maximum likelihood. The proposed method can be applied to any correlated non-Gaussian data including time series, space-time problems, and longitudinal data.

Edge Correction for Exact Tests on Nearest-Neighbor Contingency Tables for Testing Spatial Segregation

❖Elvan Ceyhan, Koc University, Department of Mathematics, Rumelifeneri Yolu Sariyer, Istanbul, 34450 Turkey, *elceyhan@ku.edu.tr*

Key Words: Association, complete spatial randomness, edge effects, exact inference, buffer zone correction, toroidal correction

Nearest neighbor contingency tables (NNCTs) are used in testing spatial patterns of segregation or association in various fields. For general contingency tables, exact tests are conservative in testing independence, but we have shown that the most conservative versions of the exact tests are appropriate for NNCTs for small samples (i.e., total number of points < 100). The null patterns are either complete spatial randomness (CSR) or random labeling (RL). The CSR model assumes that the study region is unbounded for the analyzed pattern, which is not the case in practice. In this article, we discuss the edge (or boundary) effects on various exact tests for NNCTs under the CSR case. We demonstrate that buffer zone edge correction can severely affect the results of these tests, while the toroidal correction has a mild effect on them. The right-sided exact test is robust to the edge effects.

113 Methods for and Analyses of Disease Surveillance Data ●

Section on Statistics in Epidemiology, Section on Health Policy Statistics

Monday, July 30, 8:30 am–10:20 am

Power Study of a Semiparametric Cluster Detection Method

❖Shihua Wen, University of Maryland, Mathematics Building, College Park, MD 20742, *wen@math.umd.edu*; Benjamin Kedem, University of Maryland

Key Words: semi-parametric density ratio model, false discovery rate, q-value, cluster detection, scan statistics, Combined data

A semi-parametric density ratio method which borrows strength from two or more samples can be applied to moving windows of variable size in cluster detection. The method requires neither the prior knowledge of the underlying distribution nor the number of cases before scanning. In this paper, the semi-parametric cluster detection procedure combined with controlling of the false discovery rate (FDR) for multiple testing is studied. It is shown by simulations that for binary data, using Kulldorff's Northeastern benchmark data, the semi-parametric method and Kulldorff's method performs similarly. When the data are not binary, the semi-parametric methodology still works in many cases, but Kulldorff's method requires to choose the correct scan statistic in order to achieve power comparable to that achieved by the semi-parametric method.

The Relationship Between the Recurrence Interval and Time-to-Signal Properties of Surveillance Schemes

❖Shannon Fraker, Virginia Polytechnic Institute and State University, 362 Old Cedarfield Drive, Blacksburg, VA 24060, *sfraker@vt.edu*; William H. Woodall, Virginia Polytechnic Institute and State University; Shabnam Mousavi, The Pennsylvania State University

Key Words: Average time-to-signal, Average time-between-signals, Average signaling event length, Control charts, Recurrence Interval, Scan statistics

The recurrence interval is defined to be the number of time periods for which the expected number of false alarms in a monitoring process is one. It is typically used in public health surveillance as compared to time-to-signal measures which are used in industrial statistical process control. We compare the recurrence interval and measures based on the time-to-signal properties for the temporal monitoring case using scan statistics, several control charts and Markov dependent signaling processes. The in-control average time-between-signals and the in-control average signaling event length are introduced as performance measures that are useful when a monitoring process is not reset to its initial state after a signal. We show that the recurrence interval is limited in its applicability and often fails to summarize important information about the performance of the monitoring process.

Using Administrative Data To Improve Sample Design and Estimation: Assaults Requiring Emergency Department Visits in New York City

❖Kevin Konty, New York City Department of Health and Mental Hygiene, 133 concord Street, 3a, Brooklyn, NY 11201, *konty@yahoo.com*; Catherine Stayton, New York City Department of Health and Mental Hygiene; Jingsong Lu, New York City Department of Health and Mental Hygiene

Key Words: Survey design, Disease Surveillance, Administrative Data, Syndromic Surveillance, Data quality

The NYC Department of Health's Injury Surveillance System (ISS) reviews hospital emergency department (ED) visits to estimate two quantities: the number of assaults requiring emergency room visits, and the proportion of those assaults perpetrated by a relative or partner of the victim. These estimates are used to monitor trends and evaluate and target public health initiatives. Currently, ISS reviews all ED visits for 4 weeks per year at 22 hospitals. Multipliers are used to estimate yearly assaults and the relationship of perpetrator to victim is determined by analyzing the assault data as a simple random sample. This paper describes an effort to recast the system into a probability sampling framework. We develop a novel use of syndromic surveillance ED data to improve estimation, including establishing measures of size, quantifying clustering effects, and correcting for poor coverage.

Peak-Detection in Online Influenza Monitoring

❖Eva Andersson, Göteborg University, PO box 660, Statistical Research Unit, Göteborg, SE40530 Sweden, *eva.andersson@statistics.gu.se*; Marianne Frisén, Göteborg University; David Bock, Astra Zeneca

Key Words: surveillance, on-line detection, influenza, likelihood ratio, peak detection

A non-parametric likelihood-ratio method (SRnpPeak) for peak detection is applied to Swedish influenza data. These data vary a lot between the years, with regard to the peak time, the peak height and the shape of the peak. The SRnpPeak method is derived from the optimality results of Shiryaev (1963). The method is non-parametric with regard to the shape of the peak, using previous results on ordered restricted regression. The behavior of the method is evaluated, both regarding the ability to detect peaks with different shapes, but also the effect of a larger variance than the one specified in the method. The results show that an influenza season with a low peak takes longer to detect (the delay of the alarm is longer). If the variance of the observed process is larger than the specified variance, the delay of an alarm is shorter, but the predictive value of an alarm is also lower.

Variability of Serfling's Estimator for Excess Mortality

❖Al Ozonoff, Boston University, 715 Albany Street T4E, Boston, MA 02118, *aozonoff@bu.edu*; Xiaopeng Miao, Boston College

Key Words: Influenza, Surveillance, Serfling's method, Excess mortality, Confidence interval

When considering influenza surveillance data, a primary measure of epidemic severity at the population level is excess mortality. Serfling's method to estimate excess mortality of influenza is a two-step process: (1) establish seasonal baseline by fitting a periodic regression model to a time series of recent mortality data; and (2) for future observations, calculate aggregated positive deviation above predicted baseline. Although excess mortality is a primary outcome in many epidemiological studies of influenza data, the variability of its estimator has rarely been studied, discussed, or accounted for analytically. We use simulations to quantify the variability of Serfling's estimator, and compare our results to theoretically derived values. We then apply these methods to historical influenza mortality data from the United States.

A Robust Regression Model for Estimating Influenza-Associated Deaths in 122 Cities' Mortality Data

❖Po-Yung Cheng, Centers for Disease Control and Prevention, 1600 Clifton Road MS A32, Atlanta, GA 30333, *pcheng@cdc.gov*; William W. Thompson, Centers for Disease Control and Prevention; David K. Shay, Centers for Disease Control and Prevention

Key Words: robust regression, influenza, death

CDC uses a robust regression model (Andrews) to forecast the baseline for the percentage of all deaths attributed to pneumonia and influenza (P&I) deaths using the 122 Cities Mortality Reporting System. Influenza-associated death estimates were obtained using the Andrews weight function, a Serfling linear regression model and two other weight functions (Bisquare and Huber). For the Andrews model, an annual average of 1,383 (95% CI: 947 - 1,804) P&I deaths were associated with influenza. Using the percentage of the standard error over the parameter estimates, the results indicated that the Andrews model was the best fitting model. Using these results to project national estimates, the Andrews model estimated an annual average of 4,534 (95% CI: 3,101 - 5,922) P&I deaths associated with influenza. The robust regression model provides a timely estimate of deaths associated with influenza.

Evaluation of Multiple Imputation in a Vaccine Immunogenicity Trial

❖Elizabeth Zell, Centers for Disease Control and Prevention, 3166 Bolero Dr, Atlanta, GA 30341, *ezell@cdc.gov*; Michela Baccini, University of Florence; Constantine E. Frangakis, Johns Hopkins University; Fan Li, Harvard Medical School; Fabrizia Mealli, University of Florence; Brian D. Plikaytis, Centers for Disease Control and Prevention; Charles E. Rose, Centers for Disease Control and Prevention; Donald B. Rubin, Harvard University

Key Words: Missing Data, Clinical Trials, Multiple Imputation

In randomized vaccine trials that require repeated visits, it is inevitable that some people will miss some visits. Data should be collected despite these protocol violations, but it is difficult to obtain immunogenicity data (antibody levels) or reactogenicity data (side effects) for missed visits. Some sort of imputation is required if ITT analyses are to be done using all randomized subjects. Thus it is desirable to impute the missing data. Multiple imputation has been shown to be valid under a broad range of circumstances. CDC's current Anthrax Vaccine Trial presents challenges because of the large number of measurements per person (~2000) and the limited number of subjects in each treatment arm (~200). Here we present preliminary results from a large-scale evaluation of the validity of the multiple imputations in this trial, which were created using a new state of the art procedure.

114 Advances in Latent Class and Multivariate Modeling of Survey Data ● ✪

Section on Survey Research Methods, Section on Statistics and Marketing

Monday, July 30, 8:30 am–10:20 am

Structural Equation Models for Item Parceling

❖Joseph Olsen, Brigham Young University, 578E 4380 N, Provo, UT 84604, *joseph_olsen@byu.edu*

Key Words: structural equation modeling, latent variables, item parceling, subscales, testlets

Questionnaires and tests often use many items to measure a given domain or trait. In these cases, simpler structural equation models can be developed by using parcel scores (sums or means of item subsets-subscales, testlets, or item bundles) instead of the original items as the observed indicators of the latent variables. This paper uses hierarchical factor models and confirmatory component models to outline procedures for jointly modeling parcels and their items. Using real data examples, we demonstrate the specification of item-level models which produce the same fit, estimates, and standard errors as corresponding parcel-level models. We also describe methods for testing how well the observed associations between items in different parcels can be explained by the correlation between their parcel means.

Using Factor Analysis and Cronbach's Alpha To Ascertain Relationships Between Questions of a Dietary Behavior Questionnaire

❖Eric Grau, Mathematica Policy Research, Inc., 600 Alexander Park, Princeton, NJ 08543, *egrau@mathematica-mpr.com*

Key Words: nutrition, dietary behavior, factor analysis, cronbach's alpha

Obesity and other dietary problems make it necessary to have a better understanding of dietary behavior and more effective nutrition education. A dietary behavior questionnaire was developed to measure outcomes of nutrition education as part of an effort to develop a standardized, flexible data collection tool. This questionnaire, which was separated into modules according to food groups, was field tested by Mathematica for internal consistency of responses to survey questions and the performance characteristics of individual and sets of questions. The field test data analysis identified questions that performed well and should be retained and some that performed poorly and should be either dropped or need further study. In this paper, we discuss the use of factor analysis and Cronbach's alpha to decipher the internal consistency of and relationships between questions within modules.

Randomization-Based Inference About Latent Variables in Classical Test Theory for a Clustered Sample

❖Tiandong Li, Westat, 1650 Research Blvd, Rockville, MD 20850, *tiandongli@westat.com*; Robert Mislevy, University of Maryland

Key Words: Multiple imputation, NAEP, Complex sample design, Latent constructs

Mislevy has applied "multiple imputation" techniques to Large Scale Assessment, such as NAEP, to make inference to latent constructs, which cannot be observed directly. However, in these applications, the influence of the complex sample design on the resulting inference has not been fully considered and a fixed effect model defining priors based on design and demographic variables was used in some applications. In this study, a random effect model is derived from Classical Test Theory. We focus on the estimates of variance components of a latent construct under a clustered sample design. Variation within PSUs is modeled as random effects in inferences. This approach is illustrated with a simulation study.

Hierarchical Linear Modeling for Complex Survey Data with Unequal Probability of Selection Using HLM 6

❖Frank Jenkins, Westat, 1650 Research Blvd, Rockville, MD 20850, *frankjenkins@westat.com*; Hyunshik Lee, Westat; Pam Broene, Westat

Key Words: Multilevel modeling, HLM 6.0, Growth curve analysis, Simulation study

Many educational studies concern children's growth over multiwaves of data collected via a multistage survey design. To predict growth 2- or 3-level hierarchical linear models (HLMs) can be devised employing the HLM 6 program. While standard theory asserts that all (e.g. 4) levels of the survey must be included for unbiased estimation, only 2 or 3 levels can actually be used in HLM 6. This study will build on previous research to determine bias and precision of regression and variance estimates when only part of the survey design is included in the HLM analysis. The analysis of simulated data will be more realistic than that used in previous research incorporating the following factors: Sampling weights with varying informativeness and efficiency, 3 to 5 waves of data, a range of group and strata sizes. Robust SEs using all stages of the survey will be compared to those from HLM 6.

A Methodology to Fit Hierarchical Logistic Models for Data from Complex Sampling Designs

❖Prabhu Bhagavatheeswaran, Bristol-Myers Squibb Company, 6123 Town Walk Ave, Hamden, CT 06518, *pbhagava@smu.edu*; Ian Harris, Southern Methodist University

Key Words: Hierarchical, Logistic, Method of moments, variance components

There is an increasing interest in making inferences about the relationships among the variables investigated in a survey, as opposed to simply describing the phenomenon. Hierarchical structure of data gathering, clustering, stratification and unequal selection probabilities are taken into consideration while using linear models to make inferences on model parameters. When the response variable is binary, hierarchical logistic models are more appropriate. A modified estimation procedure based on method of moments, that is computationally less intensive , is proposed here for estimating variance components in a hierarchical logistic models .Performance of this new estimator is compared with popular estimation procedures like PQL , L-6 and GH-20. The estimator is adjusted to incorporate sampling weights that arise in modeling survey data.

Post-Stratification with Optimized Effective Base: Linear and Nonlinear Ridge Regression Approach

❖Stan Lipovetsky, GfK Custom Research North America, 8401 Golden Valley Rd, Minneapolis, MN 55427, *stan.lipovetsky@gfk.com*

Key Words: Sample balance, post-stratification, ridge regression, nonlinear optimization

Post-stratification, or sample balancing, or raking, is widely utilized in survey research to weighting a sample data for a better correspondence to Census or other known population quotas. Cross-tables of counts are mostly used in the Deming-Stephan iterative proportional fitting to find the weights for adjusting data to known margins. A bi-criteria objective for finding weights with minimum variance yields a solution with maximum effective sample size. This model can be expressed as a ridge regression, which is applied to the original data, without its collapsing to cross-tables. Linear and nonlinear parameterization models are studied. The explicit regression solution allows to study the weighting analytically, that helps to interpret and improve the sample balance results.

Robustness of Latent Class Measurement Error Models

❖Brian Meekins, Bureau of Labor Statistics, 2 Massachusetts Ave, NE, Washington, DC 20212, *meekins.brian@bls.gov*; Daniell Toth, Bureau of Labor Statistics

Key Words: Markov latent class analysis, simulation, measurement error model

The technique of latent class analysis relies on a number of model assumptions which might be violated by the underlying process being investigated. This study is to determine the reliability of the analysis done on four stage Markov Latent Class models containing the classification of individuals in one of two indicator categories. The estimation is done using the EM algorithm on simulated data under specified model assumptions where those assumptions are violated to varying degrees.

115 Nonresponse and Attrition Bias in Survey Data ●

Section on Survey Research Methods

Monday, July 30, 8:30 am–10:20 am

An Analysis of Nonresponse Bias Resulting from Nonresolution of Telephone Numbers, Eligibility Screener Nonresponse, and Interview Nonresponse for the National Immunization Survey

❖Benjamin Skalland, NORC at the University of Chicago, 55 East Monroe Street, Suite 4800, Chicago, IL 60603, *skalland-benjamin@norc.org*; Robert Montgomery, NORC at the University of Chicago; Phillip Smith, Centers for Disease Control and Prevention

Key Words: Unit Non-Response, National Immunization Survey

The National Immunization Survey is a nationwide, list-assisted RDD survey conducted by NORC on behalf of the Centers of Disease Control and Prevention. It consists of two phases: a telephone component that seeks to identify households with children aged 19 to 35 months, collect socio-demographic information, and obtain consent to contact the immunization providers for those children; and a provider component wherein questionnaires are mailed to each child's providers seeking immunization data about the child, which are used to estimate vaccination coverage rates. To contribute to the estimates, then, a case must pass through five stages: determination of residential status, eligibility screening, the household interview, consent to contact providers, and provider response. This paper explores the effect of non-response at the first three stages on estimates of vaccination coverage.

Assessing the Effect of Government Frame Refinement on Collecting Establishment Data for the National Compensation Survey

❖Glenn Springer, Bureau of Labor Statistics, 2 Mass Ave NE, Rm 3160, Washington, DC 20212, *springer_g@bls.gov*

Key Words: sampling frame, nonresponse error, establishment survey

Nonresponse error is a common but undesirable feature of a survey. Survey practitioners use various techniques to reduce non-response error. In the National Compensation Survey (NCS), the state and local government frame is developed using administrative files maintained by the states for unemployment insurance (UI). In this paper, we explore the extent to which a frame refinement process, used to reconfigure UI reporting units so that they conform to how payroll records are kept in an establishment, improved the ability to collect data for the establishment as sampled, thereby improving response as well as reducing respondent burden and the need for post-collection sample-weight adjustment. An overview of the refinement process, the system tool developed for refinement, and several approaches to measuring the impact of frame refinement on non-response will be explored.

A Comparison of Level of Effort and Benchmarking Approaches for Nonresponse Bias Analysis of an RDD Survey

❖Daifeng Han, Westat, 1650 Research Blvd, Rockville, MD 20850, *daifenghan@westat.com*; David Cantor, Westat

Key Words: response rate, refusal conversion, nonresponse adjustment

One method for studying nonresponse bias is to analyze estimates by level of effort needed to complete an interview. Level of effort (LOE) is typically measured using the number of times needed to contact a respondent and/or whether the respondent had initially refused to do the interview. LOE analyses are dependent on the assumption that nonrespondents are similar to those who respond with more effort. A second method to assess nonresponse bias is to compare estimates to an external benchmark. In this paper we compare the results of a nonresponse bias analysis for an RDD survey using a LOE approach to one that compares estimates to an external benchmark of a survey with a higher response rate. Important differences are found in the results between the two approaches. One conclusion is that the assumptions behind the level of effort model may not be correct for this survey.

The Impact of Refusal Conversion on Survey Response and Error

❖ Martin Barron, NORC at the University of Chicago, NORC, 55 E Monroe Ste 1840, Chicago, IL 60603, *barron-martin@norc.org*; Karen Wooten, Centers for Disease Control and Prevention

Key Words: National Immunization Survey, Refusal Conversion, Survey Response Rates, Survey Error

Refusals are an unfortunate but inevitable part of survey research. Refusals contribute to survey error, reduce response rates and increase costs. But how many conversion attempts are appropriate? At what point does the increased cost and respondent burden outweigh the reduction in survey error? This paper explores this issue using data from the National Immunization Survey (NIS). The NIS—a nationwide, list-assisted RDD survey conducted by NORC for the Centers for Disease Control and Prevention—monitors the vaccination rates of children between the ages of 19 and 35 months. Each year, the NIS conducts interviews with approximately 29,000 households across the United States. The NIS allows up to three verbal refusals and three hang-ups without an explicit verbal refusal. Using NIS data from 2006, we model the impact of conversion attempts on survey error, cost, and response rates.

Designing Longitudinal Studies of Mobile Populations Such as Military Populations

❖ Shelton Jones, RTI International, PO Box 12194, 3040 Cornwallis Rd, Research Triangle Park, NC 27709, *smj@rti.org*; Jeniffer Iriondo-Perez, RTI International; Cynthia Bland, RTI International; Robert Bray, RTI International; Janice Brown, RTI International; Michael Pemberton, RTI International

Key Words: longitudinal studies, mobile populations, military populations, sample design

This discussion summarizes an exhaustive literature review of research focused on designing sample surveys to accommodate repeated measurements on mobile populations. Attention is given to survey design features that lead to improvements in survey response rates from the baseline of data collection through subsequent periods of data collection. Although the authors restricted their literature review primarily to studies involving military populations, similar survey methodology may be adapted to studying other populations as well. In many instances, recommendations and solutions for improving response rates will be presented to address these challenges in surveying mobile populations. Our discussion highlights design features of many military studies and important design issues and challenges typically experienced.

The 2006 American Community Survey Content Test: Addressing Measurement Errors by Pretesting, a Split-Panel Test and Reinterview

❖ Wendy Hicks, Westat, 1650 Research Blvd, Rockville, MD 20850, *wendyhicks@westat.com*; Thomas Chesnut, U.S. Census Bureau; Leah Ericson, Carnegie Mellon University; Alan Pe[illegible], U.S. Census Bureau; David Raglin, U.S. Census Bureau

Key Words: question development, split panel test, data quality

In March 2006 the U.S. Census Bureau completed the 2006 American Community Survey (ACS) Content Test. This test represents the first test of new and modified content since the ACS reached full implementation. The test included four stages: identification of questions for testing, content development and pretesting, a split-panel field test including a reinterview, and evaluation. This paper follows four individual questions with slightly different data quality issues (rooms, educational attainment, military status and health insurance) through these stages. The discussion for each question will focus on the identified error for a question designated for testing, how pretesting assessed the error, and the question design used to address the error in the split panel test. Lastly, the paper will discuss how well the tested versions addressed the errors as measured by reinterview.

116 Estimating Mortality and Migration ● ✪

Social Statistics Section

Monday, July 30, 8:30 am–10:20 am

Scepticism About The Lancet Surveys on Iraqi Mortality

❖ David Kane, Kane Capital Management, 30 Washington Street, Newton, MA 02458, *david@kanecap.com*

Key Words: Iraq, mortality, replication

Roberts et al. (2004) and Burnham et al. (2006) estimate excess mortality caused by the invasion and occupation of Iraq. Despite intensive media coverage and extensive discussion among statisticians about the technical details and merits of these studies, the authors refused to provide the underlying data and computer code that would allow outside researchers to confirm their results. How, then, can anyone know that those results are accurate? I argue that the authors should follow the replication standard, as outlined in King (1995), and provide their data and code, while maintaining respondent confidentiality. To encourage data sharing in other controversial settings, the American Statistical Association should promulgate the following Guideline: “The results of published studies which fail to meet the replication standard should be disregarded.”

Estimating Conflict-Related Mortality in Timor-Leste, 1974-1999: A Comparative Review of Demographic and Statistical Estimation Methods

❖ Romesh Silva, Human Rights Data Analysis Group, 2232 Piedmont Ave, Berkeley, CA 94708, *romesh.s@benetech.org*; Patrick Ball, The Benetech Initiative

Key Words: human rights, mortality, multiple systems estimation

Scientifically-defensible answers about the patterns and overall magnitude of conflict-related mortality can depoliticize arguments about the past. Yet, such analysis is often difficult due to limited existing data and the considerable challenges of collecting data in unsettled situations. This paper draws on a multiple data sources (retrospective survey data, administrative data and testimony data) and estimates conflict-related mortality in Timor-Leste between 1974 and 1999. Using these data, we apply survey-based estimates, multiple-systems estimation and indirect demographic estimation. Estimates of conflict-related mortality are presented along with discussion about strengths and weaknesses of the different methods. Particular focus is given to potential bias in mortality estimates derived from retrospective surveys due to positive correlation between mortality and sibship size.

Life Table Forecasting with the Gompertz Distribution

❖ Peter Pflaumer, FH Kempten, Bahnhofstr 61, Kempten, 87435 Germany, *peter.pflaumer@fh-kempten.de*

Key Words: Mortality Forecasting, Population Forecasting, Demographic Forecasting, Regression

First, this paper investigates the properties of the Gompertz distribution and the relationships of their constants. Then the use of Gompertz’ law to describe mortality is discussed with male and female period life table data of the United States between 1900 and 2000. For this purpose a model incorporating time trends has been formulated with age, time and the prod-

uct of age and time as independent variables and the force of mortality as the dependent variable. The parameters of the model are estimated using the least squares method. Important life table parameters like expectation of life, modal age, Keyfitz' entropy or old age dependency ratios are forecast up to the year 2100 and compared with recent mortality forecasts of the Social Security Administration. Our model predicts similar results for the male population, but lower mortality for the female population.

Methods of Estimating Internal Migration in the Undocumented Immigrant Population of the United States

❖ Christopher J. Campbell, U.S. Department of Homeland Security, 800 K Street NW, South Tower Suite 1000, Washington, DC 20536, *christopher.campbell@dhs.gov*; Michael D. Hoefer, U.S. Department of Homeland Security; Nancy F. Rytina, U.S. Department of Homeland Security

Key Words: Department of Homeland Security, Office of Immigration Statistics, Undocumented Immigrant Estimation, American Community Survey, Internal Migration

The Office of Immigration Statistics within the Department of Homeland Security (DHS) annually produces estimates of the undocumented immigrant population. DHS utilizes a "residual" method whereby estimates of the current legally resident population are subtracted from estimates of the foreign-born population from the American Community Survey (ACS). Estimates are produced by year of entry, country of birth, and state of residence. However, data on state of residence for the legally resident are only available at the time legally residency is obtained. Hence, the legally resident population should be migrated over time to other states within the United States to yield state estimates of the legally resident comparable to similar estimates obtained from the ACS. In this paper, we compare alternative methodologies for estimating internal migration of the legally resident population.

An Integrated Approach to Address Immigration Data Quality Problems: Sampling and Archival Research

❖ MacReadie Barr, U.S. Department of Homeland Security, 800 K St NW, South Tower Suite 1000, Washington, DC 20536, *macreadie.barr@dhs.gov*

Key Words: Department of Homeland Security, Office of Immigration Statistics, Data Quality, Sampling, Administrative Records, Archival Research

Automated immigration systems of the Department of Homeland Security have been increasingly reporting incomplete data for key variables. These variables provide important decision support for immigration policymakers regarding the human capital of immigrants, their entry into the United States, and their transition from temporary to permanent status. This study attempts to determine if the incomplete data are attributable to the applicant or the application processing. I use simple random sampling of archived paper files to determine if immigrants provide the information in question. Fieldwork was conducted at DHS's repository of immigrant records: a cave in Missouri, housing 45 million files. Results suggest that the data in question do exist. Moreover, the demonstrated sampling approach allows access to needed immigration data that cannot be obtained as efficiently elsewhere.

Status Upgrading: Immigrant Naturalization and Occupational Change Revisited

❖ Derekh Cornwell, Office of Immigration Statistics, 800 K Street NW South Tower Suite 1000, Washington, DC 20532, *derekh.cornwell@dhs.gov*; Nancy F. Rytina, U.S. Department of Homeland Security; Gary Huang, Office of Immigration Statistics

Key Words: Office of Immigration Statistics, Immigrants, Naturalization, Occupational Change, Administrative Records, Department of Homeland Security

U.S. immigration policy selects foreign nationals for legal permanent resident (LPR) status based on three criteria: family relationships, employment, and humanitarian concerns. Very little is known, however, about the status changes of LPRs relative to their category of admission due to the absence of comprehensive longitudinal data. Previous research (Jasso and Rosenzweig) has only examined naturalization and occupational outcomes for single LPR cohorts. This paper expands on this previous work by analyzing the naturalization and occupational changes for multiple LPR cohorts. Specifically, we use logistic regression to examine the likelihood of naturalization and occupational change through 2006 for immigrants obtaining LPR status between 1975 and 1995. Data for the analysis are obtained from the Department of Homeland Security.

Internal Migration of Refugee Adjustments of Status: An Application of Logistic Modeling

❖ Kelly Jefferys, U.S. Department of Homeland Security, Office of Immigration Statistics, 800 K Street NW, 10th Floor Ste 1000 South Tower, Washington, DC 20536, *kelly.jefferys@dhs.gov*

Key Words: refugee, migration, Office of Immigration Statistics, logistic, DHS

Thousands of refugees enter the United States each year, and most eventually adjust to legal permanent resident (LPR) status. Previous research has examined the initial settlement patterns of refugees in the United States (Singer and Wilson, 2006), but little is known about their internal migration patterns after arrival due to a lack of publicly available post-arrival refugee data. This study will address this research gap by matching data on refugee arrivals from the U.S. Department of State with data on refugee adjustments of status from the U.S. Department of Homeland Security, from 2001 to 2006. Logistic regression will be used to analyze the probability of internal migration among refugee adjusters of status between time of arrival and adjustment of status. The results should further our understanding of refugee settlement patterns and improve immigrant population estimates.

117 Advanced Modeling and Imaging Endpoints in Applied Oncology Statistics ● ✪

Committee on Applied Statisticians, ENAR, Section on Statistics in Epidemiology, Biometrics Section, Biopharmaceutical Section

Monday, July 30, 10:30 am–12:20 pm

Absolute Risk Models: Applications and Validation

❖ Mitchell H. Gail, National Cancer Institute, 6120 Executive Blvd, EPS 8032, Bethesda, MD 20892-7244, *gailm@mail.nih.gov*

Key Words: crude risk, pure risk, breast cancer, calibration, discriminatory accuracy, loss function

Absolute risk is the crude probability of developing a given disease in a defined age interval in the presence of competing risks. Models of absolute risk for breast cancer have been used to design trials to prevent disease, to counsel women, and to aid in deciding whether to take tamoxifen to prevent breast cancer by weighing risks and benefits of the intervention. Models for other diseases have also been developed. In this talk, I will review some of the strategies used to develop models of absolute risk, illustrate applications, and discuss approaches to model validation tailored to intended applications.

Imaging as Biomarker in Clinical Trials

❖Constantine Gatsonis, Brown University, Center for Statistical Sciences, Box G-S121-7, Providence, RI 02912, *gatsonis@stat.brown.edu*

Key Words: diagnostic tests, biomarkers, clinical trials, PET, oncology

The potential utility of quantitative imaging results as markers for response to therapy and disease progression has recently attracted significant attention. A growing body of evidence suggests that quantitative measurements obtained by imaging modalities such as Positron Emission Tomography and Dynamic-Contrast MRI can be used to predict the likelihood of response to therapy and subsequent patient outcomes. Imaging findings, alone and in combination with other markers, may then be used to make decisions about the type and course of therapy, to assess response to therapy, and to monitor patients for potential relapse. In this presentation we will discuss the statistical aspects of studies designed to evaluate the role of quantitative imaging in clinical decision making and disease management as well as in defining intermediate endpoints in clinical trials of therapy.

Some Statistical Issues with DCE-MRI Reproducibility Studies in Oncology

❖William L. Mietlowski, Novartis Pharmaceuticals, 1 Health Plaza, East Hanover, NJ 07936, *william.mietlowski@novartis.com*; Yuhui Ma, University of Medicine and Dentistry of New Jersey; Theodore C. Pellas, Novartis Pharmaceuticals; Ching-Ray Yu, Rutgers University

Key Words: DCE-MRI, reproducibility, repeatability, distribution-free, tolerance intervals

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has been used to a tool to evaluate the effect of anti-angiogenic therapies in cancer. A workshop report proposing standards for DCE-MRI (Leach et al 2005 British J Cancer) recommends incorporation of two baselines (DCE-MRI reproducibility studies) to estimate individual patient reproducibility. This degree of reproducibility is often used to determine the degree of change necessary for a patient to be classified as a DCE-MRI responder. Two methods have been used to estimate individual patient reproducibility in the imaging literature. Galbraith et al (2002 NMR Biomedicine) proposed a procedure based on repeatability while Evelhoch et al (2004 Clinical Cancer Research) proposed using the intra-patient CV. Statistical issues with these approaches will be discussed and compared with a distribution-free tolerance limit approach.

118 Recent Advances In Functional Data Analysis and Nonparametric Estimation ● ✪

Section on Physical and Engineering Sciences, Biometrics Section, Section on Nonparametric Statistics, Section on Statisticians in Defense and National Security

Monday, July 30, 10:30 am–12:20 pm

A Multiresolution Approach to Time and Image Warping

❖Bernard W. Silverman, St. Peter's College, New Inn Hall Street, Oxford OX1 2DL, United Kingdom, *bernard.silverman@spc.ox.ac.uk*

Warping functions are an intrinsic ingredient of Functional Data Analysis, because observed functions or image often need to be registered, or warped, to fit well to one another. Current approaches to the estimation of warping functions all work by modeling the function in terms of the displacements of particular points in time or space. This is not the natural approach, because warping functions are members of the group of transformations of time or space onto itself, rather than just functions. The approach I shall describe and discuss models warping functions as the composition of warplets, which are localized deformations of the time axis or image space. The order in which the warplets are composed matters, but this is something easily handled by reversible jump MCMC.

Penalized Splines: Asymptotics with Increasing Number of Knots and Equivalent Kernels

❖David Ruppert, Cornell University, School of OR & IE 225 Rhodes Hall, Ithaca, NY 14853, *davidr@orie.cornell.edu*; Yingxing Li, Cornell University

Key Words: equivalent kernel, binning

The asymptotic behavior of univariate B-splines estimated with a difference penalty is studied. Penalized splines behave like kernel estimators with "equivalent" kernels depending upon the order of the penalty. The number of knots is assumed to converge to infinity. The asymptotic distribution of the penalized spline estimate is Gaussian with simple expressions for the asymptotic mean and variance. Providing that it is fast enough, the rate at which the number of knots converges to infinity does not affect the asymptotic distribution. The optimal rate of convergence of the penalty parameter is given. Penalized splines are not in general design-adaptive. Bias decomposes into modeling and smoothing bias due to spline approximation and penalization, respectively. In our framework, the first is asymptotically negligible and the second is controlled by the penalty parameter.

Bayesian Curve-Fitting and Functional Data Analysis in Neurophysiology

❖Robert E. Kass, Carnegie Mellon University, Department of Statistics, Pittsburgh, PA 15213, *kass@stat.cmu.edu*

One of the most important techniques in learning about the functioning of the brain has involved examining neuronal activity in laboratory animals under varying experimental conditions. Neural information is represented and communicated through series of action potentials, or spike trains, and the central scientific issue in many studies concerns the physiological significance that should be attached to a particular neuron firing pattern in a particular part of the brain. We have formalized specific scientific questions in terms of point process intensity functions, and have used Bayesian methods to fit the point process models to neuronal data. I will very briefly outline some of the substantive problems we are examining and will discuss in some detail our applications involving BARS (Bayesian Adaptive Regression Splines), an approach to generalized nonparametric regression.

119 Collection and Assessment of Safety Data in a New Drug Development Program ● ✪

Biopharmaceutical Section, Committee on Applied Statisticians, ENAR, Section on Risk Analysis, Section on Health Policy Statistics, Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

A Triage Approach to Evaluating Safety Incorporating Frequentist and Bayesian Methods

❖A. Lawrence Gould, Merck & Co., Inc., UG1D-88, West Point, PA 19486, *goulda@merck.com*

Key Words: Adverse Event, Multiplicity, Order Statistics, Planning

The analysis of safety data is as important as the analysis of efficacy data. The most appropriate analysis strategy for an adverse event depends on whether it was identified a priori, as not identified a priori but not 'rare', and as not identified a priori and 'rare.' Multiplicity affects the interpretation of analyses because many adverse events often will be reported. Hypothesis testing will be of limited help for adverse events not identified a priori because the many hypotheses are generated by the data. Quantifying the degree of risk will be more helpful, especially using Bayesian methods. This presentation describes some general considerations in planning for safety evaluation, presents some ways to summarize data using confidence or credible intervals, describes a Bayesian approach to interpreting the outcomes, and suggests a simple graphical way to address multiplicity.

Assessment of Safety Data in a New Drug Development Program: Some Points to Consider

❖Weichung J. Shih, University of Medicine and Dentistry of New Jersey, School of Public Health, 16 Ambrose Valley Lane, Department of Biostatistics, Piscataway, NJ 08854, *shihwj@umdnj.edu*; Hui Quan, sanofi-aventis

Key Words: Intention-To-Treat, Treatment Emergent Adverse Event, Data Monitoring, off-drug period

Assessment of safety information in a new drug development program should be equally, if not more, important as the efficacy assessment. However, when designing a clinical trial, especially in the later phases (IIB and III), including sample size calculation and methods of evaluation, attention is usually paid more on the efficacy more than the safety side. Although a later phase study is usually regarded as a confirmatory trial by the regulatory authority, we argue that for safety analyses, it is more appropriate to emphasize their exploratory nature. In this talk we will highlight several essential points to consider along this general line.

Safety First: Managing Patients' Safety and Mitigating Products' Risks

❖ Christy Chuang-Stein, Pfizer Inc., 2800 Plymouth Road, Ann Arbor, MI 48105, *christy.j.chuang-stein@pfizer.com*

Protecting patient's safety and ensuring that benefit outweighs risk are the guiding principles in selecting treatments for patients and conducting medical experimentation. These principles were solidified in the 1948 Nuremberg Code and reinforced in the 1964 Declaration of Helsinki. The 1997 Belmont Report further reminded researchers of the need to treat subjects in an ethical manner by respecting their decisions and protecting them from harm. The above mandates monitoring safety in clinical trials and managing product safety. Increasingly, risk management plans are employed to mange marketed products (e.g., thalidomide, Lotronex and Accutane) whose usage could only be justified in a subpopulation. In this presentation, we will review the role safety evaluation plays in product assessment and the role such assessment occupies in various Prescription Drug User Fee Acts.

120 Statistics in Genomics and Proteomics ●

Biometrics Section, WNAR, Biopharmaceutical Section

Monday, July 30, 10:30 am–12:20 pm

A Method for the Detection of Alternative Splicing from Exon Array Data

❖ Terence Speed, University of California, Berkeley, 367 Evans Hall, #3860, Berkeley, CA 94720, *terry@stat.berkeley.edu*

Key Words: exon microarrays, alternative splicing

Analyses of EST data show that alternative splicing is much more widespread than once thought. The advent of exon and tiling microarrays means that researchers now have the capacity to experimentally measure alternative splicing on a genome-wide scale. New methods are needed to analyze the data from these arrays. We present a method, FIRMA (Finding Isoforms using Robust Multichip Analysis), for detecting alternative splicing in exon array data. FIRMA has been developed for Affymetrix exon arrays, but could in principle be extended to other exon arrays, or to tiling array data. We have evaluated the method using simulated data, and have also applied it to a considerable body of data from human tissues, including tumors. Some of the successes and failures of our approach will be described in the talk.

Statistical Analysis of Histone Acetylation

❖ Ping Ma, University of Illinois, 725 S Wright St, Department of Statistics, Champaign, IL 61820, *pingma@uiuc.edu*

Key Words: variable selection, epigenetics, gene expression, ChIP-chip, chromatin, motif

A comprehensive understanding of the regulatory role of histone acetylation is difficult because there exist a large number of different histone acetylation patterns and their effects are confounded by other factors such as the transcription factor binding motif information. Using the yeast as a model system, we explore an integral approach to analyze "epi-genetic" effects of histone acetylations.

Quantifying Protein: Reverse-Phase Protein Arrays

❖ Keith Baggerly, The University of Texas M.D. Anderson Cancer Center, 1100 Holcombe Blvd, Unit 237, Houston, TX 77030-3906, *kabagg@mdanderson.org*

Key Words: proteomics, bioinformatics, ELISA, immunology, preprocessing

Reverse-Phase Protein Arrays (RPPAs, aka protein lysate arrays, tissue lysate arrays, or lysate arrays) are recently developed tools for measuring protein expression levels in large numbers of samples. These assays are for the most part massively parallelized versions of enzyme-linked immunosorbent assays (ELISAs). In their massive parallelization, these assays are similar to cDNA microarrays (for mRNA) and CGH assays (for DNA). However, while those assays make thousands of measurements on a single sample ("forward-phase"), RPPAs measure one thing on hundreds of samples ("reverse-phase"). In this talk, we will attempt to place RPPAs in the broader context of other protein assays. Given this background, we will then describe issues we have encountered in modeling this data, and describe some of the tools we have developed for this purpose.

Statistical Design of Microarrays and Multiple Testing of Gene Expression Data

❖ Jason C. Hsu, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *Hsu.1@osu.edu*

Key Words: Microarrays, design of experiment, gene expresions, pharmacogenomics, closed/partition tests, permutation tests

Use of microarrays for pharmacogenomics should meet accepted statistical standards such as sensitivity and specificity requirements. We will first give recommendations for statistically designing microarrays and sample hybridization toward reproducible results, designs that have been implemented in a collaborative cancer recurrence prognosis project. Then, multiple testings ANALYSIS issues of gene expression levels will also be discussed. These issues include the definition of Type I error rate (for a single hypothesis), control of the False Discovery Rate (FDR) versus generalized Familywise Error Rate (gFWER), conditions for shortcutting closed/partition tests to step-down and step-up methods, and a subtle condition required for the validity of permutation tests.

121 Optimal Transformations

IMS, General Methodology

Monday, July 30, 10:30 am–12:20 pm

A Mean-Matching, Variance-Stabilizing Transform Approach to Nonparametric Regression in Exponential Families

❖Tony Cai, University of Pennsylvania, The Wharton School, 3730 Walnut Street, 400 Jon M. Huntsman Hall, Philadelphia, PA 19104-6340, *tcai@wharton.upenn.edu*

Key Words: adaptive estimation, block thresholding, exponential families, nonparametric regression, variance stabilizing transform, wavelets

Most of the nonparametric regression theory is developed for the case of additive Gaussian noise. In such a setting many smoothing techniques including wavelet thresholding methods have been developed and shown to be highly adaptive. In this talk we consider nonparametric regression in exponential families which include, for example, Poisson regression, binomial regression, and Gamma regression. We take the approach of using a mean-matching variance stabilizing transform to convert the problem into a standard homoskedastic Gaussian regression problem. A wavelet block thresholding method is then used to construct the final estimator of the regression function. The procedure is easily implementable. Both numerical and theoretical properties of the estimator are investigated. In particular the estimators are shown to be adaptively rate-optimal over a range of Besov Spaces.

On Optimal Smoothing Parameter Choice in Deconvolution Problems

Peter Hall, The University of Melbourne; ❖Aurore Delaigle, University of Bristol, Department of Mathematics, University Walk, Bristol, BS8 1TW United Kingdom, *aurore.delaigle@gmail.com*

Key Words: deconvolution, density estimation, smoothing parameter, regression

The SIMEX method for deconvolution enjoys a range of attractive features, among them the fact it needs little by way of tuning. However, it also suffers drawbacks, including inconsistency. For reasons such as this, and although SIMEX has enthusiasts and adherents, it is not as widely used as it might be. We suggest that SIMEX be employed for smoothing-parameter choice in other deconvolution methods (based on, for example, kernels, orthogonal series or ridging), rather than for constructing estimators itself. In this context SIMEX has very broad application, not just to different estimator types but to different deconvolution problems.

122 Medallion Lecture II

IMS, General Methodology

Monday, July 30, 10:30 am–12:20 pm

Understanding Ecological Communities

❖Claudia Neuhauser, The University of Minnesota, 100Ecology Building, 1987 Upper Buford Circle, St. Paul, MN 55108, *neuha001@umn.edu*

Key Words: interacting particle systems, ecology

Ecological communities are assemblages of populations of different species (e.g., plants, animals, microbes). Community ecology is concerned with understanding the processes that govern diversity, distribution, and abundance of species that compose an ecological community. The standard framework of mathematical models in community ecology is ordinary differential equations. This framework assumes that species are well-mixed and stochastic effects can be neglected. We present insights from work on spatially explicit and stochastic Markov processes on how local interactions and stochastic dynamics affect the outcome of community interactions, such as competition, mutualism, or parasitism.

123 Remembering Frederick Mosteller: 1916–2006

Memorial, Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Monday, July 30, 10:30 am–12:20 pm

Frederick Mosteller: A Statistical Life

❖Stephen Fienberg, Carnegie Mellon University, Department of Statistics, School of Computer Science, Pittsburgh, PA 15213, *fienberg@stat.cmu.edu*

Born in West Virginia on December 24, 1916, educated at Carnegie Tech and Princeton, Frederick Mosteller went on to a distinguished statistical career as a faculty member at Harvard University. I will present a brief biographical sketch of this remarkable statistician.

Fred Mosteller as Researcher

❖David C. Hoaglin, Abt Associates, Inc., 55 Wheeler Street, Cambridge, MA 02138-1192, *dave_hoaglin@abtassoc.com*

Frederick Mosteller was one of the most prolific and wide-ranging statistical researchers of the 20th century. In this presentation I describe some of his major contributions to statistics and to science and public policy more broadly, as well as his approaches to research.

Fred as Educator

❖Judith M. Tanur, SUNY Stoney Brook, PO Box 280, Montauk, NY 11954-0202, *jtanur@notes.cc.sunysb.edu*

Key Words: Mosteller, teaching

Statistical education has now taken on multiple facets and spawned journals and diverse collections of texts and other materials. Frederick Mosteller was a pioneer in many of the early efforts that helped set the stage for modern approaches and thinking about statistical education. In this presentation I will offer some reminiscences and examples of his contributions.

124 Recent Developments in Bayesian Methods in Data Mining/Machine Learning

Section on Bayesian Statistical Science

Monday, July 30, 10:30 am–12:20 pm

Stochastic Block Models of Mixed Membership

❖Edoardo M. Airoldi, Princeton University, Carl Icahn Laboratory, Lewis-Sigler Institute for Integrative Genomics, Princeton, NJ 08544, *eairoldi@princeton.edu*; David Blei, Princeton University; Stephen

Fienberg, Carnegie Mellon University; Eric Xing, Carnegie Mellon University

Key Words: Hierarchical Bayes, Latent variables, Mean-field approximation, Statistical network analysis, Protein interaction networks, Social networks

Observations consisting of measurements on pairs of objects arise in a variety of biological settings, with collections of author-recipient email, and in social networks. Analyses of such data typically aim at clustering the objects of study, or situating them in a low dimensional space, and at estimating relational structures among the clusters. For example, given protein interaction networks we want to estimate the memberships of individual proteins to stable protein complexes (i.e., clusters of proteins), how stable protein complexes interact with one another, and how many there are. In this talk we introduce stochastic block models of mixed membership, which support such integrated data analyses within a hierarchical Bayesian framework. A variational scheme for fast, approximate inference is presented. The methodology is demonstrated on social and protein interaction networks.

Simultaneous Regression Shrinkage, Variable Selection, and Supervised Clustering

❖Howard D. Bondell, North Carolina State University, Department of Statistics, CB 8203, Raleigh, NC 27695, *bondell@stat.ncsu.edu*

Key Words: correlation, penalization, regression, shrinkage, supervised clustering, variable selection

A new penalization technique is proposed to simultaneously select variables and perform supervised clustering. The form of the penalty function accomplishes variable selection by shrinking some coefficients to exactly zero. Additionally, this penalty yields exact equality of some coefficients, encouraging correlated predictors with similar effects on the response to form predictive clusters represented by a single coefficient. Pre-specification of the predictive clusters is not needed, thus performing the supervised clustering task within the estimation. This penalized likelihood estimator can also be viewed as the posterior mode for a particular choice of prior distribution. The procedure can be used in both regression and classification problems and compares favorably with existing approaches while yielding the added grouping information not given by typical procedures.

The Lasso with Attribute Partition Search

❖Suhrid Balakrishnan, Rutgers University, c/o DIMACS, Rutgers Univ., 96 Frelinghuysen Road, Piscataway, NJ 08854, *suhrid@cs.rutgers.edu*; David Madigan, Rutgers University

Key Words: Group Lasso, Partial Exchangeability, Time Series, Fused Lasso

Regression and classification problems involving ordered attributes (for example where some input patterns are a set of samples from time series variables) arise in application domains like finance and epidemiology. In such cases, identifying and building models involving predictive runs of the attributes leads to highly interpretable models that may also be very accurate. We present an approach to build such models using a variant of the Group Lasso (Yuan and Lin, 2006).

125 A Sampling Revolution? List versus Area Frames for Probability Sampling ● ✪

Section on Survey Research Methods

Monday, July 30, 10:30 am–12:20 pm

Predicting the Relative Quality of Alternative Sampling Frames

❖Colm O'Muircheartaigh, NORC at the University of Chicago, 1155 East 60th Street, Chicago, IL 60637, *colm@norc.uchicago.edu*; Edward English, NORC at the University of Chicago; Stephanie Eckman, NORC at the University of Chicago

Key Words: USPS lists, sampling frames, Tailored samples, frame quality

The paper presents a systematic review of the relative efficacy of traditional listing and the USPS address list as sampling frames for national probability samples of households. NORC and ISR collaborated to compare these two national area-probability sampling frames for household surveys. We conducted this comparison in an ongoing survey operation which combines the current wave of the HRS with the first wave of NSHAP. Since 2000, survey samplers have been exploring the potential of the USPS address lists to serve as a sampling frame for probability samples from the general population. We report the relative coverage properties of the two frames, as well as predictors of the coverage and performance of the USPS frame. The research provides insight into the coverage and cost/benefit trade-offs that researchers can expect from traditionally listed frames and USPS address databases.

Improving Coverage of Residential Address Lists in Multistage Area Samples

❖Sylvia Dohrmann, Westat, , *dohrmas1@westat.com*; Daifeng Han, Westat; Leyla Mohadjer, Westat

Key Words: Area samples, USPS, address lists

Address lists originating from the United States Postal Service (USPS) can be used as area sampling frames in place of on-site enumerations of dwelling units. While it has become clear that purchased USPS lists are less costly than the process of enumeration, it is less clear whether these lists are adequate as substitutes for them. We will evaluate the coverage of purchased lists for a selection of primary sampling units, differing in size and composition. We will do this by forming second stage units (SSUs) using Census geography, and determining which of the USPS addresses fall into those SSUs using on-site enumeration. We will then focus on ways to improve the coverage of these lists and discuss the impact of discovering missed units on field operations. We will also examine practical ways of assigning selection probabilities to the missed units.

Comparing the Coverage of a Household Sampling Frame Based on Mailing Addresses to a Frame Based on Field Enumeration

❖Vincent G. Iannacchione, RTI International, 701 13th ST NW Suite 750, Washington, DC 20005, *vince@rti.org*; Joseph McMichael, RTI International; James R. Chromy, RTI International; David Cunningham, RTI International; Katherine Morton, RTI International; James Cajka, RTI International; Ross Curry, RTI International

Cost savings and flexibility are primary advantages of using mailing addresses instead of field enumeration as a sampling frame for household (HH) surveys. A sampling frame based on mailing addresses can be developed in weeks instead of the months usually required for field enumeration. The question is whether the advantages of mailing addresses are accompanied by a decrease in the coverage of the household population. Our research is based on a probability sample of 50 Census Blocks (CBs) in North Carolina. Within each CB, we constructed a frame of HHs based on mailing addresses and a frame based on field enumeration. We used GPS technology to match the HHs from each frame, and not presume that either approach is the gold" standard. We estimate the prevalence of HHs in the two frames for urban and rural areas, areas without home delivery of mail, and areas with gated communities.

126 Multiplicity Today: Perspectives on Multiple Testing ● ✪

General Methodology, Section on Statistics in Epidemiology, WNAR

Monday, July 30, 10:30 am–12:20 pm

Political Aspects of Multiple Testing: A Controversy

❖S. Stanley Young, National Institute of Statistical Sciences, 19 T.W. Alexander Drive, P.O. Box 14006, Research Triangle Park, NC 27709-4006, *young@niss.org*

Key Words: Multiple testing, False positives

Some multiple testing mistakes are due to ignorance (how often are you asked to re-examine the data to see if something can be found?), but others are intentional, planned and well-understood policy (over $1B of grant/tax money flows to institutions with reproducibility problems revolving around multiple testing). Statisticians need to understand both ignorance and devious policy. It serves neither society nor our profession to ignore multiple testing controversies. At a minimum we need to protect the integrity of our profession. We present an egregious treatment of multiple testing and review a light-hearted multiple testing paper. We present survey of journal editors on multiple testing. How can you improve the situation? When you suspect a false positive due to multiple testing, write a letter to the editor. Ask for the data set and if it is not available, protest to the journal.

Empirical Bayes Methods for Multiple Testing

❖Debashis Ghosh, University of Michigan, 1420 Washington Heights, Department of Biostatistics, Ann Arbor, MI 48109/2029, *ghoshd@umich.edu*

Key Words: Decision Theory, Selection, Gene Expression, Microarrays, Massively Univariate, Simultaneous Inference

Scientific areas such as genomics and neuroimaging have generated high-dimensional data that has necessitated the consideration of multiple comparisons procedures. In addition, for many situations, the choice of null hypothesis is not appropriate. In this talk, we discuss how the Empirical Bayes approach pioneered by Robbins is very appropriate to the scientific setting at hand. We summarize the use of Empirical Bayes methods for selecting hypotheses in the multiple testing problem. We then discuss how one can construct estimators and confidence intervals for this situation as well. The issue of the appropriate null hypothesis, raised by Efron (2004), is mentioned here; accommodations of the methodology to various "empirical null" hypotheses from the literature are shown. Finally, we show how to extend the procedure to incorporate prior knowledge.

Hochberg's Step-Up Method: Cutting Corners off Holm's Step-Down Method

❖Yifan Huang, University of South Florida, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Dr USF, MRC205, Tampa, FL 33612, *huangy@moffitt.usf.edu*; Jason C. Hsu, The Ohio State University

Key Words: Hochberg's method, Holm's method, Multiple comparisons, Step-up test, Step-down test, Partition testing

Holm's and Hochberg's methods are popular multiple tests. They are viewed as step-down and step-up versions of the Bonferroni test. We will give the insight that both are special cases of partition testing. The difference is that, while Holm's method tests each partition hypothesis using the largest order statistic, setting a critical value based on the Bonferroni inequality, Hochberg's method tests each partition hypothesis using all the order statistics, setting a series of critical values based on Simes' inequality. Geometrically, Hochberg's step-up method 'cuts corners' off the acceptance regions of Holm's step-down method by making assumption on the joint distribution of the test statistics. As can be expected, partition testing making use of the joint distribution of the test statistics is more powerful than partition testing using probabilistic inequalities.

Multiple Tests of Association with Biological Annotation Metadata

❖Sunduz Keles, University of Wisconsin-Madison, Madison, WI , *keles@stat.wisc.edu*; Sandrine Dudoit, University of California, Berkeley; Mark J. van der Laan, University of California, Berkeley

Key Words: Multiple testing, Type-I error rates, Resampling-based approaches, Gene set enrichment analysis, Biological meta data, Gene ontology

We propose a general and formal statistical framework for the multiple tests of associations between known fixed features of a genome and unknown parameters of the distribution of variable features of this genome in a population of interest. The known fixed gene-annotation profiles may concern Gene Ontology (GO) annotation, regulation by particular transcription factors, and pathway membership. The gene-parameter profiles may be, for example, regression coefficients relating genome-wide transcript levels or DNA copy numbers to clinical outcomes. A rigorous formulation of the inference question allows us to apply the multiple testing methodology developed in Dudoit and van der Laan (2006) to control a broad class of Type I error rates, in testing problems with general data generating distributions. The proposed methods are illustrated using the ALL dataset of Chiaretti et al. (2004).

127 Spatial Surveillance for Adverse Environmental Health Outcomes ● ✪

Section on Statistics and the Environment, Section on Bayesian Statistical Science, Section on Statisticians in Defense and National Security, ENAR, Section on Health Policy Statistics, Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

Optimal Methods in Surveillance

❖Marianne Frisén, Göteborg University, Statistical Research Unit, Box 640, Göteborg, SE40530, Sweden, *marianne.frisen@statistics.gu.se*

Key Words: Surveillance, Monitoring, Changepoint

Timely detections of environmental health hazards are important both at an individual level and to society. Statistical methods are necessary to separate important changes in the process from stochastic variation. New types of systems to analyze the data and meet the demand of surveillance have been developed recently. The statistical methods suitable for this, differ from the standard hypothesis testing methods. Good properties are quick detection, few false alarms and robustness. An overview of optimality criteria and methods will be given. Different ways to evaluate surveillance systems and to analyze data are demonstrated by freeware. Swedish data on biodiversity, radiation and infectious diseases will be used to illustrate methods. References can be found at *www.statistics.gu.se/surveillance.*

Bayesian Modeling and Surveillance for Adverse MRDD Outcomes Associated with Soil Chemical Exposures

❖Ji-In Kim, University of South Carolina, Dept. of Epidemiology & Biostatistics, Arnold School of Public Health, Columbia, SC 29208, *jiinkim@gwm.sc.edu*; Andrew B. Lawson, University of South Carolina

Key Words: environmental exposure, logistic, spatial, MRDD, clustering interpolation

The relation between early childhood development and maternal exposures to environmental chemicals during pregnancy is an important issue when considering the residential exposure risk. In this study we examine a range of modeling methods where we have geo-coded residential addresses for mothers during the different months of pregnancy and mental retardation and development delay (MRDD) outcome measures for the babies for a Medicaid population in South Carolina. We also have available measures of soil chemistry (e.g. total microtox EC50) on a network of sites. Our modeling involves interpolation methods for spatially-referenced measures to locations of residence that vary with time. We also develop a logistic spatial model for the MRDD outcome and clustering in that outcome which can be time-dependent or designed to be a function of the cumulative exposure over all addresses resided in.

Assessing Spatial Performance of Surveillance Systems

❖Lance Waller, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, *lwaller@sph.emory.edu*; Yuemei Wang, Emory University; Carol Lin, Bristol-Myers Squibb Company

Key Words: spatial statistics, surveillance, system performance, sensitivity, specificity

Many disease outbreak surveillance systems are designed based on concepts from statistical process control and assess performance based on temporal summaries such as the expected time until false or true alarms. While these important temporal aspects of performance play key roles in assessing performance, spatial aspects of system performance also merit attention. More specifically, most environmental surveillance systems often monitor health outcomes or exposures across a defined geographic space, but this space is not homogeneous with respect to local sample size and local uncertainty. We explore aspects of system performance addressing spatial summaries of statistical power (specificity) and sensitivity (1 - false alarm rates), and illustrate the importance of evaluating both when and where events of interest are likely to occur in determining system performance.

128 *JASA,* Theory and Methods Invited Session

JASA, Theory and Methods, General Methodology

Monday, July 30, 10:30 am–12:20 pm

Implementation of Estimating-Function–Based Inference Procedures with MCMC Samplers

❖Lee-Jen Wei, Harvard University, Department of Biostatistics, 655 Huntington Avenue, Room 445, Bldg. II, Boston, MA 02115, *wei@hsph.harvard.edu*; Lu Tian, Northwestern University; Jun S. Liu, Harvard University

Key Words: Bootstrap, Median regression, Metropolis algorithm, Normal approximation, Resampling, Survival analysis

Under a semiparametric or nonparametric setting, inferences about the unknown parameter are often made based on a non-smooth estimating function. Resampling methods are quite handy for obtaining good approximations to the distribution of the consistent estimator when the estimating equation and its resampled counterparts are not difficult to solve numerically. In this paper, we propose a simple, flexible procedure which provides such approximations via the standard Markov chain Monte Carlo sampler without solving any equations. More generally the procedure may locate all possible roots of the estimating equation and provides an approximation to the distribution of each root. We illustrate our proposal extensively with three examples. The performance of the new procedure is also examined comprehensively via a simulation study.

129 SAMSI: A Five-Year Retrospective and Five-Year Prospective ● ✪

Statistical and Applied Mathematical Sciences Institute

Monday, July 30, 10:30 am–12:20 pm

SAMSI: A Five-Year Retrospective and Five-Year Prospective

❖Peter Bickel, University of California, Berkeley, *bickel@stat.berkeley.edu*; ❖Sallie Keller-McNulty, Rice University, Brown School of Engineering, Houston, TX 77251-1892, *sallie@rice.edu*; ❖Alan Karr, National Institute of Statistical Sciences, PO Box 14006, Research Triangle Park, NC 27709-4006, *karr@niss.org*

Key Words: SAMSI, interdisciplinary, future research, institute, infrastructure, remote access

In July 2007, the Statistical and Applied Mathematical Sciences Institute will be five year's old. During these five years SAMSI will have conducted 16 research programs spanning an enormous range of the statistical sciences and their interactions with other fields. The panel will first address the major themes and 'lessons learned' from this extensive SAMSI research, as well as changes in remote access technology (which will be illustrated) that have made SAMSI more widely accessible to the national (and international) statistics community. Next, the panel will turn to the future, and initiate a wide-ranging discussion as to the directions that SAMSI should pursue over its next five years, in terms of both research topics and mode of operation and infrastructure. The audience is very much encouraged to participate in this discussion; please come to the session with ideas for the future.

130 Issues and Solutions to Planning and Implementing an Adaptive Design in Practice ●

ENAR, Biometrics Section, Biopharmaceutical Section, WNAR

Monday, July 30, 10:30 am–12:20 pm

Issues and Solutions to Planning and Implementing an Adaptive Design in Practice

❖Michael Krams, Wyeth, 500 Arcola Road A-5113, Collegeville, PA 19426, *kramsm@wyeth.com*; ❖Paul Gallo, Novartis Pharmaceuticals, 419/2114, One Health Plaza, East Hanover, NJ 07936, *paul.gallo@novartis.com*; ❖Gernot Wassmer, The University of Cologne, Institute for Medical Statistics, Informatics, and Epidemiology, Kerpener Str. 62, 50937 Cologne, International 50937 Germany, *gernot.wassmer@uni-koeln.de*; ❖Jerald S. Schindler, Merck & Co., Inc., 126 E. Lincoln Avenue, Rahway, NJ 07065, *Jerald_Schindler@merck.com*; ❖Christopher S. Coffey, The University of Alabama at Birmingham, Department of Biostatistics, 1665

University Blvd., Birmingham, AL 35294, *ccoffey@ms.soph.uab.edu*

Key Words: Adaptive Designs, Flexible Designs, Data Monitoring Committee, planning, NIH

Adaptive trial designs can be a valuable addition to the drug development toolkit, but more objective evaluation based on experience with their use in practice is needed. Statistical, logistical, and procedural issues need to be addressed. For example, some key statistical issues are control of the type I error rate and unbiased estimation. Some logistical issues are the timely collection of quality data, and the need for the integration of data capture, drug supply management, and communication between patient/investigator and the randomization center. A key procedural issue is the processes by which interim data is reviewed and decisions are made and implemented in ways to ensure that the integrity of the trial are not compromised. In this session, 5 panel members across academia and industry will discuss issues and practical solutions based on their experiences.

131 It's Only Data, What's the Big Deal?: The Significance of Using Multiyear Estimates from the American Community Survey ● ✪

Social Statistics Section, Section on Survey Research Methods

Monday, July 30, 10:30 am–12:20 pm

Multiple ACS Estimates: Pick a Number, Any Number!

❖Linda Gage, California Department of Finance, Demographic Research Unit, 915 L Street, Sacramento, CA 95814, *linda.gage@dof.ca.gov*

Key Words: American Community Survey, ACS, Census, estimates

The American Community Survey will produce annual one-, three-, and five-year estimates for large communities; annual three- and five-year estimates for others; and only annual five-year estimates for the smallest communities. The first nation-wide five-year estimates will begin in 2010. A new research data set produced by the Census Bureau provides data from several ACS sites where data has been collected since 1999. This allows the evaluation of the multi-year estimates. This paper focuses on the usefulness and accuracy of ACS data for San Francisco and Tulare counties in California. It considers issues that will arise in assisting data users: how stable is the time series of annual ACS data, which of the multiple estimates for a single jurisdiction and single year should be used, and how does the ACS compare with available administrative data sources?

Multiyear ACS Results for Multnomah County, Oregon: Preliminary Observations

❖George C. Hough, Portland State University, PO BOX 751 - PRC, Portland, OR 97307-0751, *houghg@pdx.edu*; David A. Swanson, University of Mississippi

Key Words: American Community Survey, stability, User interpretation burden, Multnomah County, Oregon

In 2005, the American Community (ACS) was implemented nation-wide for households. ACS is a Census Bureau product designed to provide annually updated demographic, social, economic, and housing indicators for both large and small geographic areas within the United States. Operational plans for Census 2010 call for ACS to replace the decennial census long-form (Census LF). Because of sample size and design issues, questions have arisen about the stability of annual level ACS estimates, particularly for areas with populations less than 20,000. This paper examines the stability of annual estimates for ACS results in Multnomah, County, Oregon and selected subareas, including the City of Portland. With an eye toward maintaining a level of "User Interpretation Burden" equivalent to that of the decennial census long form, we discuss the implications of our findings for ACS users.

The Utility of Five-Year Averages from the American Community Survey for Bronx Neighborhoods

❖Joseph Salvo, NYC Department of City Planning, 22 Reade Street, New York, NY 10007, *jsalvo@planning.nyc.gov*; A. Peter Lobo, NYC Department of City Planning; Adam Willett, New York City Department of City Planning

Key Words: American Community Survey, Census, small area data, data quality, administrative data, survey data

The capacity of the American Community Survey (ACS) to replace the decennial census long form is based on the validity and reliability of these data at the census tract level. This research builds on an earlier analysis of ACS test data for the Bronx that utilized special three-year averages for neighborhoods, which were census tract aggregates. The availability of census tract data for the Bronx, which are five-year averages, provides an ideal opportunity to ask about the quality and usefulness of these data, which will become available for all census tracts in 2010. This study evaluates the reliability of these census tract attributes against those provided by the 2000 Census. Further, we attempt to gauge the usefulness of these estimates by comparing the ACS five-year averages to administrative data at the census tract level for selected population, housing and economic items.

132 Applications of Visualization for Web 2.0 ● ✪

Section on Statistical Graphics

Monday, July 30, 10:30 am–12:20 pm

An AJAX Web 2.0 Geospatial Visualization Framework

❖Stephen Eick, University of Illinois at Chicago, 725 SEO, M/C 249, 851 S. Morgan St., Chicago, IL 60607-7045, *eick@cs.uic.edu*

Key Words: AJAX, thin client, geospatial visualization, Web 2.0, statistical graphics, visualization components

Recently a new class of applications has emerged that uses AJAX and other web programming techniques to provide a rich user experience in a web browser. This class of applications is being called Web 2.0 and includes Google Maps and Google Suggest. To experiment with this approach, we have developed a Web 2.0 thin client collaborative visualization framework called GeoBoostT that uses Scalable Vector Graphics and AJAX to provide a rich user experience built around collaboration. Our framework includes geospatial maps, standard business charts, node and link displays, and custom visual displays. All of our visualization components run in standard web browsers and provide rich interaction and collaboration.

Statistical Graphics with Element Control in the Browser

❖Sven Knudsen, Insightful Corporation, Puglundvej 40, Varde, 6800 Denmark, *sknudsen@insightful.com*; Michael O'Connell, Insightful Corporation

We describe a new graphics model with element control in a browser environment. The graph properties are specified as name-value pairs based on the property element syntax used in XAML. The graph container is based on the S-PLUS Graphlet, lightweight applets rendered as XML in the browser. Text, line, symbol, fill, legend, axes, area and polygon are controlled as elements; text area controllers are also included. We illustrate the model with a palette of graphics types and patterns covering analysis, review and reporting of clinical safety and efficacy data.

Using Web 2.0 for Statistical Software

❖Webster West, Texas A&M University, Department of Statistics, College Station, TX 77843-3143, *websterwest@yahoo.com*

Key Words: Web 2.0, software, Web-based

Web 2.0 refers to next generation Web-based applications. These applications typically run within a web browser without the need for any plug ins. Many of these new web technologies provide a rich graphical environment which is ideal for analyzing data. A survey of existing Web 2.0 technologies will be provided along with a discussion of how these technologies may be applied to develop statistical software.

Statistical Graphics for Collaborative Environments

❖Daniel Rope, SPSS Inc., 2000 N 14th St Suite 320, Arlington, VA 22201, *drope@spss.com*

New technology has enabled a truly collaborative environment on the web where ideas are formed by communities rather than individuals. Blogs, Wikis, and several other communal platforms allow anyone to contribute to any topic of their choice. This talk will examine the role that statistical graphics has begun to play for these collaborative environments and explore the tremendous possibilities as well as some of the risks.

133 When To Stop a Clinical Trial for Efficacy? ● ✪

Biopharmaceutical Section, ENAR, WNAR

Monday, July 30, 10:30 am–12:20 pm

When To Stop a Clinical Trial for Efficacy?

❖Vipin Arora, Novartis Pharmaceuticals, One Health Plaza, Ridgedale Avenue, East Windsor, NJ 07936-1016, *vipin.arora@novartis.com*

Key Words: Stopping boundaries, Clinical Trials, Interim analysis, Patient, Sponsor

Stopping boundaries are pre-specified for clinical trials with interim analyses (IA) and detailed in the Data Monitoring Committee (DMC) Charter. These boundaries should be stringent enough to protect patient and sponsor interests in demonstrating positive efficacy before normal completion of the study. O'Brien-Fleming (O-F), Haybittle-Peto (H-P), Pocock family type boundaries are common choices. Selection of stopping boundaries may impact the patient and sponsor interests due to early trial termination. Experienced DMCs are very well aware of their responsibility and their assessment is based on totality of evidence. Most commonly, H-P type boundaries are used when there are serious ethical concerns and the O-F are used when the intentions of stopping are genuine. Examples based on these two methods to safeguard interest of all stakeholders including patient and sponsor will be presented.

Overview of Available Methods for Efficacy Monitoring

❖Thomas Cook, University of Wisconsin-Madison, 209 WARF Office Building, 610 Walnut Street, Madison, WI 53726, *cook@biostat.wisc.edu*

In formulating the monitoring plan for a long term randomized trial, one is faced with a wide array of possible monitoring boundaries. While many boundaries are constructed to satisfy a mathematical optimality criterion such as small average sample number, most boundaries are chosen because they subjectively balance a set of competing interests; ethical, practical and economic. Many potential features are of interest. Boundaries maybe one or two sided. If two sided, they may be symmetric or asymmetric. One may take a flexible, alpha spending approach or a use a more traditional approach such as Haybittle-Peto. Boundaries based on simultaneous alpha/beta spending rates are also available, automatically taking futility into account. Finally, standard computational techniques are available which allow commonly used boundaries to be employed for a wide variety of statistical tests.

Crossing Monitoring Boundaries and Not Stopping

❖Janet Wittes, Statistics Collaborative, 1625 Mass Ave NW, Suite 200, Washington, DC 20036, *janet@statcollab.com*

Key Words: randomized trials, sequential analysis, DSMB

A DSMB monitoring a randomized clinical trial uses prespecified boundaries as guidelines for its actions. If the trajectory of the data crosses the "upper" boundary, the trial can stop with a declaration of "success." Although DSMBs are aware that these boundaries are only advisory, they recognize that stopping without crossing leads to an inflated Type I error rate and can therefore jeopardize the scientific credibility of the results. On the other hand, failing to recommend stopping a trial when a boundary is crossed incurs no statistical penalty. DSMBs are reluctant to continue after crossing because the boundary reflects the Board's opinion prior to the start of the trial as to the data that it would find convincing evidence of benefit. In the context of actual examples, the paper discusses reasons a DSMB may not recommend stopping a trial even after crossing a boundary for benefit.

134 Statistical Challenges for Bioassay Data ● ✪

Biopharmaceutical Section, WNAR

Monday, July 30, 10:30 am–12:20 pm

Prediction-Based Decision for Validation of (Bio)Analytical Methods Using Tolerance Intervals and Accuracy Profiles

❖Bruno Boulanger, Université de Liège, Chemin du Foriest, Braine L Alleud, B1420 Belgium, *bruno.boulanger@ulg.ac.be*; Walthère Dewe, GlaxoSmithKline; Francois Moonen, Arlenda; Philippe Hubert, University of Liege

Key Words: validation assays, tolerance interval, accuracy profile, fit-for-purpose, prediction based decision

The purpose of validation of analytical methods is to demonstrate that it is suited for its intended purpose. For quantitative analytical method the objective is to quantify unknown samples with an acceptable accuracy. A decision about the suitability of the analytical method based on prediction is proposed at validation stage: a method is declared proper for routine use if most of the future results it will generated are likely to be accurate. This is achieved by using the "fl-expectation tolerance interval" and the corresponding "accuracy profile." A "fit-for-purpose" approach also proposed to select the most efficient standard curve. Four case studies will be reported showing that predictions made using the "fl-expectation tolerance intervals" can be trusted for decision making. This way the objectives of the analytical method, the objectives of the validation phase are conciliated.

Statistical Methods for Assessing Long-Term Stability of Compounds in Biological Matrices

❖David M. Hoffman, sanofi-aventis, 1041 Route 202206, Mailcode M203A, Bridgewater, NJ 08807, *david.hoffman2@sanofi-aventis.com*; Robert Kringle, sanofi-aventis

Key Words: stability, variance components, mixed model, regression

The purpose of a long term stability study is to demonstrate that clinical samples are stable over the maximum period of sample storage expected in future clinical trials. Because these studies often last two years or longer, a particular difficulty is dealing with frequently large random month-to-month variability. Various study designs, analysis methods, and decision procedures have been proposed in the literature. We propose a new design-analysis procedure that utilizes a concurrent control sample to adjust for random month-to-month fluctuations in a mixed model regression analysis. Performance of the new procedure is evaluated and compared to other procedures by simulation and by application to historical studies.

Bioassay: Decisionmaking in the Face of Uncertainty

❖ Timothy Schofield, Merck & Co., Inc., Merck PO Box 4, Building 37C305, West Point, PA 19486, *tim_schofield@merck.com*

Key Words: Pharmaceuticals, Bioassay, Risk

Bioassays are utilized in the pharmaceutical industry to monitor and control therapeutic proteins and vaccines. These are important because they are usually linked to the mechanism of activity of the pharmaceutical. Bioassays are subject to variability, which translates into elevated decision risks during development and manufacture. These risks can be minimized through appropriate bioassay design, optimization, and validation. Studies and control strategies utilizing bioassays can be designed to overcome the impact of inherent bioassay variability. This presentation will highlight opportunities to minimize variability, and thereby risk. Bioassay designs such as parallel-line analysis together with randomization, as well as manufacturing and QC models, will be highlighted for mitigating risk to the customer and manufacturer alike.

Reparameterization of the Five-Parameter Logistic Dose-Response Curve

❖ Jason Liao, Merck & Co., Inc., POBox 4, West Point, PA 19486, *jason_liao@merck.com*; Rong Liu, Merck & Co., Inc.

Key Words: dose response curve, four parameter logistic, five parameter logistic, asymmetry, assay, potency

Recently, the five-parameter-logistic function (5PL) has been increasing used to model a dose-response curve. The 5PL takes into the curve asymmetry into consideration to overcome the drawback of currently widely used four-parameter-logistic (4PL) function, which could lead to wrong estimates of the important parameters, to poor characterization of the pharmacological pathways and the mechanisms, and to wrong inference of drug-receptor interactions. However, the currently used 5PL form does not have good practical interpretation of these parameters comparing to the currently used 4PL. In this talk, we will re-parameterize the 5PL to match the good parameter interpretations of the 4PL. The re-parameterized new 5PL has better practical interpretations and statistical properties.

Overcoming Practical, Statistical, and Robotic Challenges in Bioassay

❖ David Lansky, Precision Bioassay, 257 S Union St, Burlington, VT 05401, *david@precisionbioassay.com*

Good animal husbandry, statistical design, and analysis are all needed for a successful bioassay. Identifying experimental units is challenging for properly randomized assays; for the all-too-common non-randomized assay the challenges grow. Practical constraints in the laboratory often create multiple levels of blocking, serial dilution error, and design structures more complex than split-units. Relative potency is a non-linear function of the observed nonlinear response and there is often non-additivity in the variation of the response. With non-constant variance, outliers, and insufficient access to statisticians familiar with bioassay there are many opportunities for ad-hoc "solutions." By combining good designs, statistical software to drive laboratory robots, and mixed model analyses we aim to provide a comprehensive solution that facilitates doing bioassay well.

135 Innovative Methods for Analyzing Observational Studies and Clinical Studies ●

IMS, Section on Statistics in Epidemiology, ENAR

Monday, July 30, 10:30 am–12:20 pm

Robust Propensity Score Analysis for Causal Inference in Observational Studies

❖ Asheber Abebe, Auburn University, Dept of Mathematics and Statistics, 221 Parker Hall, Auburn, AL 36849, *abebeas@auburn.edu*; Joseph W. McKean, Western Michigan University; Bradley E. Huitema, Western Michigan University

Key Words: Wilcoxon, R-estimation, matching, program evaluation

The use of propensity scores for matching observations based on pre-treatment variables has seen increasing use in recent years. This method, however, is very sensitive to outlying observations. In this presentation, a robust version of propensity score matching that is based on rank-based procedures is proposed. The large and small sample performance of this new method are evaluated. Case studies using real life data sets are presented.

Propensity Score Analysis with Hierarchical Data

❖ Fan Li, Harvard Medical School, 180 Longwood Ave, Boston, MA 02115, *li@hcp.med.harvard.edu*; Alan M. Zaslavsky, Harvard University; Mary Beth Landrum, Harvard Medical School

Key Words: propensity score, hierarchical model, causal inference, health policy, Bayesian

The propensity score method is being increasingly used as a less parametric alternative to the traditional regression methods in many fields . However, propensity score was developed and have been applied in cross-sectional settings with unstructured data. Data collected in the health policy studies are typically clustered or hierarchically structured, in the sense that patients are grouped together in one or more ways. In this talk, we present results in the following aspects: (1) develop and apply valid propensity score methods for causal inference in the context of hierarchically structured data; (2) elucidate assumptions underlying methods for causal inference in the context of hierarchically structured data; and (3) develop diagnostics, which make use of the clustered structure, for testing the sensitivity of causal inferences to important violations of assumptions.

Bayesian Hidden Markov Models for Alcoholism Data

❖ Kenneth Shirley, University of Pennsylvania, 400 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, *kshirley@wharton.upenn.edu*; Dylan Small, University of Pennsylvania; Kevin G. Lynch, University of Pennsylvania

Key Words: Hidden Markov Models, Alcoholism, Longitudinal Data, Gibbs Sampler, Forward-Backward Recursion

Alcoholism data typically consists of daily drink counts for recovering alcoholics, and covariates, some of which are time-varying and are also measured on a daily basis. Hidden Markov Models do a good job of capturing different drinking behaviors over time for this type of data. This paper incorporates covariates into an HMM for alcoholism data to model relationships between covariates and the outcome. Bayesian methods are used to fit the models, and predictions from the model are compared to those from other models. The clinical interpretability of estimates from the HMM is emphasized.

Laplace's Approximation for Relative Risk Frailty Models

❖Shibao Feng, Genentech, Inc., 1 DNA Way MS 441B, South San Francisco, CA 94080-4990, *shibaof@gene.com*; Lei Nie, Georgetown University Medical Center; Robert A. Wolfe, University Renal Research and Education Association

Key Words: Multivariate survival analysis, Mixed Poisson regression models, Empirical Bayes estimates

Relative risk frailty models are used extensively in analyzing clustered and/or recurrent time-to-event data. In this paper, Laplace's approximation for integrals is applied to marginal distributions of data arising from parametric (e.g. piecewise exponential) relative risk frailty models. A full likelihood approach (Feng, Wolfe, and Port 2005) is used to estimate the parameters. Under regularity conditions, the full likelihood estimators are shown to be consistent with a rate of convergence depending on both the number of subjects and number of observations per subject. We compare the full likelihood estimators against alternative estimators using limited simulation and demonstrate the utility of the full likelihood approach by analyzing U.S. patient waiting time to deceased kidney transplant data.

Dental Arch Clustering in Orthodontics Practice

❖Johan Lim, Yonsei University, Applied Statistics, Seoul, 120-749 South Korea, *johanlim@yonsei.ac.kr*; Heon Jin Park, Inha University; SungIm Lee, DanKook University; Shin-Je Lee, Seoul National University

Key Words: clustering, dental arch

In orthodontics practice, it has an important meaning to cluster dental arch forms of human adults with normal occlusions into a few sets of patterns according to their geometric similarities. In this work, we propose a method to cluster dental arch forms of normal occlusion samples; each sample consists of 28 points measured on 2-D plane 14 points are for maxilla and the other 14 points are for mandible. Two key components of the proposed method is: First, we define the distance between two arch form as the area between two arches in 2-D plane. Second, we plot the observations on polar coordinates rather than on Cartesian coordinates due to several merits we will see in the presentation. We apply our method to clustering dental casts of 307 normal occlusion samples which are selected from 15,836 adults through a community dental health survey from the year of 1999 to 2002 in Seoul.

136 Statistics of Extremes ✪

IMS

Monday, July 30, 10:30 am–12:20 pm

Large Deviations for Point Processes Based on Stationary Sequences with Heavy Tails

❖Henrik Hult, Brown University, Division of Applied Mathematics, 182 George St, Providence, RI 02912, *henrik_hult@brown.edu*

Key Words: regular variation, point processes, large deviations, heavy tails

A stationary sequence of random variables with regularly varying tails is considered. For this sequence it is possible that large values arrive in clusters. That is, there may be many large values in a relatively short period of time. The aim is to give a detailed description of the occurrence of large values, not only the size of clusters but also within a single cluster. To do this a point process based on appropriately scaled points of the stationary sequence is constructed. A limiting measure, on the space of point measures, describes the joint limiting behavior of all the large values of the sequence. From this limiting result one can proceed and obtain the functional large deviation for its partial sums, ruin probabilities, etc. Examples include, linear processes, random coefficient ARMA processes, and solutions to stochastic recurrence equations. Joint work with G. Samorodnitsky.

Inference of a Heavy-Tailed Distribution

❖Yongcheng Qi, The University of Minnesota Duluth, Department of Mathematics & Statistics, 1117 University Drive, Duluth, MN 55812, *yqi@d.umn.edu*

Key Words: Confidence interval, Coverage probability, Empirical likelihood, Tail index estimation, Edgeworth expansion

This paper proposes some new estimators for the tail index of a heavy tailed distribution when only a few largest values are observed within blocks. These estimators are proved to be asymptotically normal under suitable conditions, and their Edgeworth expansions are obtained. Empirical likelihood method is also employed to construct confidence intervals for the tail index. The comparison for the confidence intervals based on the normal approximation and the empirical likelihood method is made in terms of coverage probability and length of the confidence intervals. The simulation study shows that the empirical likelihood method outperforms the normal approximation method.

Maximum Likelihood Estimation for Alpha-Stable Autoregressive Processes

❖Beth Andrews, Northwestern University, 2006 Sheridan Road, Evanston, IL 60208, *bandrews@northwestern.edu*; Matthew Calder, Colorado State University; Richard A. Davis, Colorado State University

Key Words: Autoregressive models, Maximum likelihood estimation, Noncausal, Non-Gaussian, Stable distributions

We consider maximum likelihood estimation for the parameters of autoregressive time series processes with non-Gaussian α-stable noise. Both causal and noncausal autoregressive models are considered. Applications for these time series models have appeared, for example, in the fields of economics and finance, signal processing, and teletraffic engineering. We give a nondegenerate limiting distribution for consistent maximum likelihood estimators of model parameters. The estimators for the autoregressive model parameters are $n^{1/\alpha}$-consistent, while the estimators for the parameters of the stable noise distribution have the traditional $n^{1/2}$ rate of convergence. The behavior of the estimators for finite samples is studied via simulation, and maximum likelihood estimation is used to fit a noncausal autoregressive model to stock market trading volume data.

On the Estimation of the Heavy-Tail Exponent in Time Series Using the Max-Spectrum

❖Stilian Stoev, University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109, *sstoev@umich.edu*; George Michailidis, The University of Michigan

Key Words: heavy-tail exponent, max-spectrum, block-maxima, moving maxima, max-stable, time series

The estimation of the tail index of distributions with heavy, Pareto-type tails is an old problem which continues to pose challenges. We propose a novel approach for estimating the tail index, based on the max self-similarity scaling of block maxima. The method exploits the increasing lack of dependence of maxima over large size blocks, which proves useful for dependent, time series data. We establish the consistency of the proposed max-spectrum estimator for certain classes of dependent time series and demonstrate its robustness to short-lived contaminations in the data. The max-spectrum estimator exhibits linear computational time and memory complexity and can be calculated in a sequential manner, that makes it particularly well suited both for massive, as well as streaming data sets. It also provides a natural time-scale perspective in the analysis of heavy-tailed time series.

Quotient Correlation: A Sample-Based Alternative to Pearson's Correlation

❖ Zhengjun Zhang, University of Wisconsin-Madison, Department of Statistics, Madison, WI 53706, *zjz@stat.wisc.edu*

Key Words: necessary condition of tail independence, nonparametric statistical coefficient, semiparametric statistical coefficient, dependence measure, gamma distribution, extreme value distribution

The quotient correlation is defined here as an alternative to Pearson's correlation that is more intuitive and flexible in cases where the tail behavior of data is important. It measures nonlinear dependence where the regular correlation coefficient is generally not applicable. One of its most useful features is a test statistic that has high power when testing nonlinear dependence in cases where the Fisher's Z-transformation test may fail to reach a right conclusion. Unlike most asymptotic test statistics, which are either normal or χ^2, this test statistic has a limiting gamma distribution (henceforth the gamma test statistic). More than the common usages of correlation, the quotient correlation can easily and intuitively be adjusted to values at tails. This adjustment generates two new concepts—the tail quotient correlation and the tail independence test statistics.

137 Applications of Objective Bayesian Analysis ● ✪

Section on Bayesian Statistical Science

Monday, July 30, 10:30 am–12:20 pm

Estimate Response Rates and Satisfaction Given Response Jointly Using Hierarchical Bayesian Modeling Approach

❖ Xiaoming Gao, Missouri Department of Conservation, 1110 South College Ave, Columbia, MO 65203, *sherry.gao@mdc.mo.gov*; Dongchu Sun, University of Missouri-Columbia; Zhuoqiong (Chong) He, University of Missouri-Columbia

Key Words: Bayesian Hierarchical model, Spatial corrrelation, nonresponse, 2-fold CAR model, conditional probabilities

Human survey is an important tool in wildlife management. Missouri Department of Conservation (MDC) conducts Hunter Attitude Survey periodically to estimate hunters' satisfaction with MDC's deer management program. In an effort to increase response, the survey was sent out three mailings, but the state wide response rate is only 55%. In this study, we're interested in estimating response rate and satisfaction rate given response at county level. We lay out it as a multinomial problem based on conditional probabilities. We propose a Bayesian hierarchical model to characterize the mailing and county effects on the response and satisfaction rates. This model also includes a 2-fold CAR model which allows us to estimate not only the spatial correlations in response rates and satisfaction rates given response among counties, but also a correlation estimate between response and satisfaction.

Objective Priors for Overdisposed Weibull Models for Breast Cancer Incidence and Survivals

❖ Luyan Dai, University of Missouri-Columbia, 29 Broadway Village Dr, Apt F, Columbia, MO 65201, *ld9n9@mizzou.edu*; Zhuoqiong (Chong) He, University of Missouri-Columbia; Dongchu Sun, University of Missouri-Columbia; Mario Schootman, Washington University in St. Louis

Key Words: Refernce prior, spatial-survival model, posterior quantiles, objective priors, breast cancer

Most of the widely used Weibull models in Bayesian survival analysis are based on subjective priors. The Bayesian estimates are often very sensitive to the choices of the priors. Some known objective priors for overdisposed Weibull models would result in improper posteriors. In this paper, we propose a spatial-survival model by incorporating reference priors to investigate geological variation on county and census tract level among breast cancer incidents in Iowa. Moreover, the estimators of unknown parameters are computed by posterior quantiles since reference priors lead to nonexistence of posterior moments in this case. Results based on subjective and objective priors are also compared.

Objective Priors for Spatially Adaptive Smoothing Splines

❖ Yu Yue, University of Missouri-Columbia, 52 Broadway Village Dr, Apt A, Columbia, MO 65201, *yytc9@mizzou.edu*

In previous work, Yue and Speckman developed a class of priors for spatially adaptive smoothing and thin-plate splines by introducing a spatial model for the precision parameter of an intrinsic Gaussian Markov random field. The adaptive priors are effective in the classic nonparametric regression problem with normal error terms (i.e., curve fitting). The computation is efficient due to the sparseness of the precision matrix. In this work, we discuss the possibility of using objective priors on the variance terms of the adaptive priors. The results are mixed. Unlike the nonadaptive case investigated by Speckman and Sun (2003), it appears that improper priors cannot be used simultaneously at all levels of the hierarchical model. However, we find conditions on proper priors in the high levels of the model to permit the use of the invariance prior for the error variance.

Objective Bayesian Analysis in Memory Study

❖ Xiaoyan Lin, University of Missouri-Columbia, Statistics Dept., 29 Broadway Village Dr, Apt F, Columbia, MO 65201, *xlzt3@mizzou.edu*; Dongchu Sun, University of Missouri-Columbia

Key Words: Objective Bayisan, Hierarchical Model, Right Haar prior, Two-level Wishart prior, Gibbs sampler

Psychologists believe memory consists two primary components, "conscious recollection" and "automatic activation." A famous approach to measure the two memory components is the Process Dissociation Procedure. Two sets of generalized linear additive models are used to model the recollection ability and automatic activation ability, respectively. Additive components are the effect of a participant and the effect of an item. Objective Bayesian analysis is conducted to estimate the effects. We put the hierarchical priors on the generalized linear models. At the first stage, two bivariate normal priors are used for the effects of participants and the effects of items, respectively. At the second stage, we use objective priors on the two covariance matrices. The simulation is done to illustrate the goodness of the model. The propriety of the posterior distribution is explored.

138 Section on Bayesian Statistics Student Paper Competition: Applications in Biostatistics ● ✪

Section on Bayesian Statistical Science
Monday, July 30, 10:30 am–12:20 pm

Multivariate Spatial-Temporal Modeling and Prediction of Speciated Fine Particles

❖ Jungsoon Choi, North Carolina State University, P219 ES King Village, 3820 Jackson St, Raleigh, NC 27607, *jchoi@ncsu.edu*; Montserrat Fuentes, North Carolina State University; Brian Reich, North Carolina State University; Jerry Davis, North Carolina State University

Key Words: multivariate spatiotemporal processes, Bayesian inference, ; linear coregionalization model, air pollution, environmental statistics

PM2.5 is an atmospheric pollutant linked to serious health problems. It is composed of sulfate, nitrate, total carbonaceous mass, ammonium, and crustal material. These components have complex spatiotemporal dependency and cross-dependency structures. It is important to understand the patterns of these components and to estimate how they might change with space and time, by spatially interpolating speciated PM2.5 to conduct spatiotemporal epidemiological studies of the association of these pollutants and adverse health effects. We introduce a multivariate spatiotemporal model for speciated PM2.5 using a linear coregionalization model and a Bayesian hierarchical framework. We also introduce a statistical framework to combine different sources of data to explain bias and measurement error. We apply our framework to speciated PM2.5 data in the US in 2004.

Antiviral Dynamic Models with Application to AIDS Studies: A Bayesian Approach

❖ Tao Lu, 14305 Wedgewood Circle, Apt 21, Tampa, FL 33613, *tlu2004@hotmail.com*

Key Words: Bayesian nonlinear mixed-effects models, McMC, longitudinal data, long-term HIV dynamics, antiretroviral drug therapy, AIDS

A virologic marker, the number of HIV RNA copies [illegible] viral load, is currently used to evaluate antiviral therapies in AIDS clinical trials. This marker can be used to assess the antiviral potency of therapies but is easily affected by drug exposures, drug resistance and other factors during the long-term treatment process. To better understanding of the factors responsible for the virologic failure, this paper develops a mechanism-based nonlinear differential equation models for characterizing long-term viral dynamics with ARV therapy. A Bayesian nonlinear mixed-effects modeling approach is investigated for estimating dynamic parameters by fitting the model to viral load data from an AIDS clinical trial. These results suggest viral dynamic parameters may play an important role in understanding HIV pathogenesis designing new treatment strategies for long-term care of AIDS patients.

Bayesian Variable Selection with Joint Modeling of Categorical and Survival Outcomes: An Application to Individualizing Chemotherapy Treatment in Advanced Colorectal Cancer

❖ Wei Chen, Karmanos Cancer Institute, Biostatistic Core, 4100 John R, Detroit, MI 48201, *chenw@karmanos.org*

Key Words: Multivariate Regression, Hierarchical Model, Latent Variable, Interaction, MCMC

Clinical studies for colorectal cancer have shown that genetic alterations lead to different responses to same treatment, despite the morphologic similarities of tumors. This article focuses on developing statistical method appropriate for individualizing treatment. The multivariate regression model with latent variables and structured variance covariance matrix considered here accounts for the correlated nature of multiple endpoints and the fact that endpoints have different statistical distributions. The mixture normal hierarchical structure incorporates a variable selection rule for interaction terms. The application to the advanced colorectal cancer study revealed the associations between multiple endpoints and certain alterations of biomarkers, demonstrating the potential of individualizing treatment based on genetic profiles.

Expert Opinion, Informative Priors, and Sensitivity Analysis for Longitudinal Binary Data with Informative Dropout

❖ Joo Yeon Lee, Food and Drug Administration, 1733 Anderson RD, Falls Church, VA 22043, *jooyeon@stat.brown.edu*; Joseph W. Hogan, Brown University

Key Words: Informative dropout, Pattern mixture model, Sensitivity analysis, Prior elicitation, Informative priors

Elicitation of expert opinions is a crucial role in fully Bayesian way of sensitivity analysis. In this paper we use a pattern mixture model, composed of transition model within pattern, for repeated binary data with nonignorable dropout and introduce parameter identification scheme which allows the analyst to explore the effects of possibly nonignorable dropout. We show how to elicit prior distributions that reflect beliefs about the distribution of missing responses using experts opinions and how different prior beliefs can affect study conclusion. Methods are illustrated using data from the OASIS study, a longitudinal clinical trial of a motivational intervention for smoking cessation in smokers participating in outpatient alcohol treatment.

A Semiparametric Joint Modeling Approach for Nonignorable Missing Data

❖ Liansheng Zhu, Pharmaceutical Product Development, Inc, 311 Castle Hayne Dr, Cary, NC 27519, *liansheng.zhu@rtp.ppdi.com*; Sujit Ghosh, North Carolina State University

Key Words: Bayesian inference, GAM, Joint modeling, Missing data

It is generally difficult to analyze incomplete data when the missingness is informative. Pattern-mixture models are commonly used in practice. Concerns include that not all parameters are identifiable and marginalization over patterns can be tricky for discrete outcomes. Zhu et al. (2007) proposed a parametric joint-modeling approach within a Bayesian framework in which patterns are defined by surrogate variables and treated as random effects. A parametric generalized linear model may not be appropriate especially when the missingness is non-ignorable. An extension to accommodate nonlinear models is proposed within the framework of generalized additive models (GAMs). Our small simulation study indicates that the joint model by using GAMs performs well when the underlying model is non-linear. We reanalyzed the CPCRA trial data and confirmed the significance of the treatment effect.

139 Nonparametric Curve Estimation ● ✪

Section on Nonparametric Statistics
Monday, July 30, 10:30 am–12:20 pm

A New Bayesian Procedure for Nonparametric Regression Estimation and Pointwise Confidence Bands

❖ Linda Zhao, University of Pennsylvania, 3730 Walnut ST, Philadelphia, 19104, *lzhao@wharton.upenn.edu*

We construct approximate Bayesian confidence intervals for the values, f(x), of a nonparametric regression function. The method uses a blockwise construction which reduces the problem to that of confidence intervals for single coordinates of a multivariate normal mean. The prior adopted for the multivariate mean problem is a modification of the Harmonic prior. In this multivariate problem the Bayesian estimates perform well although the Bayesian regions for single coordinates tend to overcover in a crucial part of the space. A frequentist modification is then suggested in the multivariate problem and the results are carried over to the original regression problem. The resulting nonparametric regression confidence bands have outstanding frequentist coverage properties over a wide range of empirical examples.

Consistent Learning Methods Are Approximately Local

❖ Yaacov Ritov, The Hebrew University of Jerusalem, Department of Statistics, Jerusalem, 91905 Israel, *yaacov@mscc.huji.ac.il*; Alon Zakai, The Hebrew University of Jerusalem

Key Words: Nonparametric, Classification, SVM

We investigate the following learning-theoretical statement: We call a learning method local if, given a training set, it produces an estimate for a particular point that depends only on close-by points from the training set. We formally define this concept in several ways and show several theoretical results relating to those definitions, in particular that any consistent method is very close to being local.

Remarks on Bandwidth Selection for Discrimination

❖ David W. Scott, Rice University, 4143 Marlowe Street, Houston, TX 77005-1953, *scottdw@rice.edu*

Key Words: kernel density estimation, classification, bandwidth, smoothing

Unlike parametric estimation, no single nonparametric estimate is optimal for different purposes. Using nonparametric density estimates in statistical discrimination, Friedman (1997) argued that optimal smoothing parameters should be much larger than for optimal density estimation itself. That is, there is a much different bias/variance tradeoff for discrimination. We revisit this thesis, borrowing on an idea of Hall and Wand (1988) using density differences.

Estimating Common Mean Curve in Interlaboratory Studies

❖ Andrew Rukhin, University of Maryland, Department of Math Stat, 1000 Hilltop Circle, Baltimore, MD 21250, *rukhin@math.umbc.edu*

Key Words: meta-analysis, statistical metrology

The primary goal of this work is to extend two methods of random effects models to situation when a consensus mean in a collaborative study is represented as a curve. These methods comprise the DerSimonian-Laird estimator, stemming from meta-analysis, and the Mandel-Paule algorithm widely used in metrological interlaboratory studies. The alternative estimators are also discussed. Two methods of assessing the uncertainty of these estimators and the resulting confidence bands are given.

Conditional Density Estimation

❖ Sam Efromovich, University of Texas at Dallas, 4213 Twilight Trail, Plano, TX 75093, *efrom@utdallas.edu*

Key Words: Adaptation, Regression, Nonparametric, Minimax

A problem of estimation of the conditional density of the response given the predictor is considered. A universal adaptive estimator is suggested which: (i) matches performance of an oracle that knows both an underlying model and an estimated conditional density; (ii) is sharp minimax over a vast class of anisotropic conditional densities; (iii) is at least rate minimax when the response is independent of the predictor and thus a conditional density becomes a univariate marginal density (performs a dimension reduction); (iv) is adaptive to an underlying design (fixed or random) of predictors. Some extensions, including multiwavelet estimators and vector-predictors, are also discussed.

140 Statistical Literacy 2007

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Monday, July 30, 10:30 am–12:20 pm

Statistics for Innumerate Journalists

❖ Stephen Doig, Arizona State University, Cronkite School of Journalism, Campus Box 871305, Tempe, AZ 85287-1305, *steve.doig@asu.edu*

Key Words: journalism, innumeracy, statistics education

By and large, journalists tend to be bright people who at some point in public school hit a bad math teacher and from then on decided “I must be a word person, not a numbers person.” Unfortunately, good journalism requires competent math skills, to the horror of many otherwise good journalists. This paper will describe strategies for teaching basic math and statistics, from descriptive to linear regression, to math-phobic reporters.

Quantitative Reasoning: An Activity-Based Course with Real Data and Relevant Issues

❖ Kay Somers, Moravian College, 1200 Main Street, Bethlehem, PA 18018, *mekbs01@moravian.edu*

Activity-based learning that uses real data and draws on real-life situations is the basis for our entry-level, general education course “Quantitative Reasoning and Informed Citizenship.” As its title suggests, the course investigates relevant applications and is designed to help students become more informed citizens and learn to use technology effectively. This talk will provide an overview of the course topics, which are organized into three general sections: Numerical Reasoning; Logical Reasoning; and Statistical Reasoning. We will discuss the variety of contexts, from sports to social and economic issues, and the types of examples, exercises, and activities that have been designed to engage students. We will describe the classroom activities and will present positive results that illustrate how students’ attitudes about and aptitude for mathematics are affected by such a course.

Success of Relative Strength Investing Within Similar Asset Classes

❖ Marshall Schield, Stir Research, 665 Glacier Club Drive Unit 6, Durango, CO 81301, *mls@stirresearch.net*; Milo Schield, Augsburg College

A mutual fund analysis of Callan data found that buying last year’ winners gave less growth than simply diversifying the portfolio each year. The mutual fund analysis implied that using relative strength to make purchases was suboptimal. However, the same study with repeated with dramatically different conclusion. Buying last years leaders gives the investor a tremendous lead. Why the difference? The difference is whether the asset classes are similar or different. The Callan table includes a bond as an asset class, along with 7 equity asset classes. The bond class is extremely influential. It is like a confounder; it distorts the data. If you do the same study using only equity classes, buying last year's leader works extremely well. This phenomenon is verified using different types of assets.

Grammar of Chance

❖Tom Burnham, Cognitive Consulting, 7550 Windsor Oaks, San Antonio, TX 78239, *tvvb@vortex4.net*; Milo Schield, Augsburg College

Key Words: statistical literacy

The grammar of chance is reviewed. Chance grammar includes odds, risk, likelihood and probability for single percentages. It also includes likely and prevalent along with other comparisons of ratios. Chance grammar is often used in everyday speech in a non-quantitative sense or if calculated from data, the data may not constitute a reasonable statistical universe. In chance statements, it is often difficult to distinguish part from whole or to distinguish the random outcome from the generating process. Examples are reviewed from a variety of sources. This paper examines grammatical devices that are sufficient to eliminate such ambiguities.

Media Stories Involving Numbers

❖Milo Schield, Augsburg College, 1767 Alameda, Roseville, MN 55113, *milo@pro-ns.net*

Key Words: statistical literacy, numeracy, quantitative literacy

Several hundred web-based media stories involving numbers have been analyzed for their statistical content. This content included the description and comparison of counts and measures, percentages and rates, chance and probability. Traditional statistical topics include sample size, confidence level, margin of error, and statistical significance. Non-traditional statistical topics include types of arithmetic association, effect size and the percentage and number of cases attributable to an exposure. Topics involving the statistical context included the type of study (experiment vs. observational, longitudinal vs. cross-sectional, controlled vs. uncontrolled) and the presence of random selection and random assignment. The prevalence of these topics is analyzed and compared. This web-accessible index may help teachers find media stories that illustrate a particular quantitative topic.

141 Statistical Phylogenetics

IMS

Monday, July 30, 10:30 am–12:30 pm

A Bayesian Model To Infer the Relationship Among Several Recombinant Sequences

❖Erik Bloomquist, University of California, Los Angeles, Department of Biostatistics, School of Public Health, Los Angeles, CA 90095, *ebloomqu@ucla.edu*; Marc A. Suchard, University of California, Los Angeles

Key Words: Phylogenetics, HIV, Bayesian, Recombination, HBV

We propose a Bayesian model that allows the relationship among several recombinants to be inferred. The model also provides dating information for the recombinant virions. The model is an extension of a previous model introduced by Minin et al. (2005). The model is highlighted by the use of an HIV example as well as a Hepatitis B example.

Inference About Gorilla Population History Using MCMC Method

❖Joungyoun Kim, University of Wisconsin-Madison, 4725 Sheboygan avenue, 337, Madison, WI 53705, *joungyoun@gmail.com*

Key Words: Phylogenetics, Genealogy, Population tree, Sequence tree, Coalescent process

The purpose of this paper is to explore the geographical distribution of genetic variation in gorilla populations in order to better understand the phylogeographic history of them. We base our inference on two types of DNA sequence data. One type is HV1, the first hyper-variable region in the mitochondrial control region. The other type is so-called Numts, DNA sequences in the nuclear chromosomes which are thought to be copies of HV1. We expect that the substitution rate in mitochondrial sequences is much higher than in nuclear genes. Therefore, our model uses different rates for each sequence type. We want to estimate (1) the phylogenetic relationships among gorilla populations; (2) times of the events when mitochondrial sequence was transferred to nuclear chromosomes. We take a Bayesian approach using MCMC methods for the computation.

Bayes Estimation of Hybridization in Species Phylogeny Based on Incongruent Gene Topologies

❖Chen Meng, University of New Mexico, 521 Spruce St SE Apt 308, Albuquerque, NM 87106, *chenmeng@unm.edu*

Key Words: phylogenetics, hybridization, coalescence, Bayesian

Evolutionary inference via phylogenetic trees has been increasingly performed on DNA sequence data from multiple genes for a set of organisms. Both incomplete lineage sorting and hybridization (gene flow between distinct species) may cause topological incongruence between species tree and gene trees. We develop a model that assumes independent evolution across genes and allows both deep coalescence and hybridization. Frequentist tests and Bayesian approaches are used to detect hybrid speciation. The Bayesian method incorporates prior knowledge about hybridization. An MCMC algorithm is implemented to estimate the proportion of the genes that are derived from each parental species under the model and to determine which factor (hybridization or coalescence) is primarily responsible for variation in individual gene histories. We apply the method to simulated and real data.

Joint Bayesian Estimation of Phylogeny and Sequence Alignment

❖Heejung Shim, University of Wisconsin-Madison, Statistics, 402 Grand Ave 5, Madison, 53705, *shim@stat.wisc.edu*

Key Words: phylogeny, sequence alignment, MCMC

Phylogeny and sequence alignment are estimated separately in the traditional techniques : first estimate a multiple sequence alignment, and then infer a phylogeny based on the sequence alignment estimated in the previous step. We develop a joint model for co-estimating phylogeny and sequence alignment. The phylogeny estimate from the joint model avoids bias due to conditioning on a single alignment and properly accounts for alignment uncertainty. We model the insertion/deletion process directly without artificial restrictions against overlapping events. Since our method does not use dynamic programming, we expect improvement in time complexity in comparison to other methods for joint phylogeny and sequence alignment estimation. We use a Bayesian approach using MCMC to estimate the posterior distribution of phylogenetic tree and a multiple sequence alignment.

Spatial Location-Dependent Substitution Models and Their Application to Phylogenetics

❖Xueliang Pan, The Ohio State University, 3536 Chowning Ct, Columbus, OH 43220, *xpan@stat.ohio-state.edu*; Dennis K. Pearl, The Ohio State University; J. Dennis Pollack, The Ohio State University

Key Words: phylogeny, Bayesian method, substitution model, crystallography

The conservation of the amino acids in homologous DNA sequences depends on a selection process that depends on functional and structural constraints. In this study, a model allowing site variation (conservation) to depend on spatial location is presented to reconstruct the phylogeny using Markov Chain Monte Carlo methods. The spatial location of each amino acid is obtained independently from crystallographic images. The model then assumes that conservation is monotonically increasing as we get closer in space to the functional core of a protein. Differences in quality of fit, in convergence and computational efficiency between this model and some other widely used models allowing rate variation are compared. We demonstrate the usefulness of the model by analyzing large data sets with sequences from 12 different enzymes which are each found in species from all Kingdoms.

142 Theory and Applications of Gene Coexpression Network Methods ● ✪

WNAR, ENAR, Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

Microarray Gene Coexpression Analysis by a Supervised Sequential Clustering Algorithm

❖Cheng Cheng, St. Jude Children's Research Hospital, Department of Biostatistics MS 768, St Jude Childrens Research Hospital, Memphis, TN 38016, *cheng.cheng@stjude.org*

Key Words: Clustering, Gene, Co-expression, Microarray, Pathway, Genomics

Most microarray gene expression analyses have focused on the analysis of each gene (probe) individually. Because genes form complex signaling and regulatory pathways in carrying out their biological functions, existence of coexpressions are plausible in even a snapshot microarray gene expression experiment, and such coexpressions may be reflected by the stochastic relationships among the genes' expression data. Detecting the coexpressions may help identify the underlying signaling/regulatory pathways involved in the biological process of interest. A clustering algorithm for this purpose, coupled with an inference procedure, will be presented in this talk. The methodology will be illustrated by real-data examples.

WGCNA: A Software for Weighted Gene Coexpression Network Analysis

❖Lin Wang, University of California, Los Angeles, 11511 Ohio Avenue, Apt 4, Los Angeles, CA 90025, *wlfll@yahoo.com*

Key Words: Weighted network, gene expression, gene co-expression

Weighted gene co-expression network approach for analyzing microarray data has been shown to be useful to elucidate the molecular basis of cancer, complex traits, and other genomic applications. Here we present a software, WGCNA, which implements gene co-expression network construction, module detection, and connectivity calculation. We show how to use the software to find biologically and statistically significant genes. This user-friendly software has a graphic interface that facilitates straightforward input of microarray and clinical trait data. For each gene it computes several intra-modular connectivity measures, which can be interpreted as fuzzy clustering. We also illustrate network screening by identifying cancer related genes and show how to compute a false discovery rate.

Network-Based Meta Analysis Methods for Microarray Data

❖Wen Lin, University of California, Los Angeles, 29 Alevera St, Irvine, CA 92618, *wenlin@ucla.edu*; Steve Horvath, University of California, Los Angeles

Key Words: network

The accumulation of microarray data poses a new statistical problem: how to combine multiple gene expression data that have been measured using different technological platforms and different measurement protocols. Toward this end, several meta-analysis techniques have been suggested in the literature. Since traditional meta analysis techniques may ignore important biological pathway, we propose a novel method that makes use of weighted gene co-expression network concepts. Specifically, we propose two types of network-based scores that allow one to rank genes based on their gene significance levels and module membership information. Using simulations and real data applications, we compare the performance of these network based meta analysis techniques to several existing approaches including a recent method based on the concept of "integrative correlation" (Parmigiani et al. 2004).

Clustering Procedures for Finding Shared Modules Across Multiple Microarray Datasets

❖Peter Langfelder, University of California, Los Angeles, Dept of Human Genetics, 695 CE Young Dr S Box 708822, Los Angeles, CA 90095-7088, *peter.langfelder@gmail.com*; Steve Horvath, University of California, Los Angeles

Key Words: Gene co-expression networks, Module networks, gene dissimilarity, hierarchical clustering

We develop a novel statistical method for finding shared modules across multiple microarray gene expression datasets. This data reduction technique employs a ``consensus'' gene dissimilarity measure (based on weighted gene co-expression networks) in conjunction with hierarchical clustering. Several suitable dissimilarity measures are presented and compared. We briefly explore their statistical properties, then apply the most appropriate method to several empirical datasets. In gene expression networks, the shared modules represent pathways that are common among the organisms or tissues under study and hence are biologically meaningful.

Connectivity, Module-Conformity, and Significance: Understanding Weighted Gene Coexpression Networks

❖Jun Dong, University of California, Los Angeles, Dept of Human Genetics, Los Angeles, CA 90095, *jundong@ucla.edu*; Steve Horvath, University of California, Los Angeles; Andy Yip, National University of Singapore

Key Words: connectivity, hub gene, clustering coe±cient, eigengene, co-expression network, microarrays

Network concepts are increasingly used in biology and genetics. Here we study the relationship of important network concepts in gene coexpression module networks based on the transcriptional response of cells to changing conditions. The main goal of this paper is to explain the meaning of intramodular connectivity and other network concepts in terms of the underlying gene expression profiles. We present theoretical and empirical results that show how intramodular connectivity and several other network concepts are related to properties of the module eigengene. Our theoretical results allow us to characterize modules where hub genes will be highly correlated with an external microarray sample trait, and are applicable to any network that is comprised of highly correlated genes (e.g., a gene coexpression module). We illustrate our results with a brain cancer coexpression network.

143 Recent Advances in Microarray Data Analysis I ● ✪

Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

Exonerating Bonferroni's Multiple Testing Procedure

❖ Alexander Gordon, The University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, *aygordon@uncc.edu*; Galina Glazko, University of Rochester; Xing Qiu, University of Rochester; Andrei Yakovlev, University of Rochester

Key Words: multiple testing, Bonferroni's procedure, stability, microarrays

The Bonferroni multiple testing procedure is commonly perceived as being overly conservative in large-scale simultaneous testing situations such as those that arise in microarray data analysis. The objective of the present study is to show that this popular belief is due to overly stringent requirements that are typically imposed on the procedure rather than to its conservative nature. To get over its notorious conservatism, we advocate using the Bonferroni selection rule as a procedure that controls the per family error rate (PFER). Its stability, as well as the ability to provide strong control of the PFER, make the Bonferroni procedure an attractive choice in microarray studies.

Statistical Framework for Integrative Analysis of Multiple Gene Expression

❖ Chien-Cheng Tseng, University of Pittsburgh, 130 Desoto Street, Pittsburgh, PA 15218, *ctseng@pitt.edu*

Key Words: meta-analysis, microarray, expression profile

With the availability of tons of expression profiles on the web, the needs of meta-analyses to enhance different types of microarray analyses are obvious. For detection of differentially expressed genes, most of the current efforts are focused on comparing and evaluating gene lists obtained from each dataset and a real sense of information integration is rarely performed. In this paper, we propose a general framework of statistical integration by a weighted averaged statistics when multiple biologically relevant data sets are available and a permutation analysis is applied to control the false discovery rate. A subset of differentially expressed genes identified in the integrative analysis are otherwise ignored in each individual analysis. We will show the advantage of combining information from multiple data sets through simulation and real data of lung and prostate cancer.

A Method of Detecting Differential Gene Expression for Cross-Species Hybridization Experiments

❖ Ying Chen, University of California, Davis, 4400 Solano Park Circle, APT 1211, Davis, CA 95616, *ychen@wald.ucdavis.edu*; Anu Chakicherla, Lawrence Livermore National Laboratory; David Rocke, University of California, Davis

Key Words: cross-species, microarray, gene set analysis, GSEA, ToTS

Cross-species microarray hybridization has been shown to be a potentially powerful tool for understanding genomics in organisms in which there is no complete genomic sequence available. However, sequence differences between two species made it difficult to detect any significant change in gene expressions. Previous literature demonstrated an empirical masking procedure to analyze such data. The fact that gene set analysis is useful to detect minimal or moderate changes in gene expressions motivated us to apply this technique in cross-species analysis. In this study, we proposed a novel knowledge-based method to analyze a real cross-species data set. We compared the results of two gene set analysis tools (GSEA and ToTS) used in this data set. Our findings showed that gene set analysis method can be used as a powerful technique for analyzing cross-species hybridization experiments.

Identifying Differentially Expressed Genes for Time-Course Microarray Data Through Functional Data Analysis

❖ Kun Chen, University of California, Davis, Department of Statistics, Davis, CA 95616, *kchen@wald.ucdavis.edu*; Jane-Ling Wang, University of California, Davis

Key Words: Time-course Gene Expression, Differentially Expressed Genes, Functional Data Analysis, Functional Principal Component, hybrid EM

In order to identify genes that show differential expression under two conditions, we develop and employ the Functional Principal Component (FPC) model to depict the dynamics of the gene trajectories. In our model, the gene expression trajectories are captured by less basis functions than the other methods, and these basis functions are estimated from the data reflecting major mode of variation in the data. The dependency structure of the gene expression over time is also considered without any parametric assumptions and estimated from all genes such that the information across the other genes can be borrowed. The parameter estimation is carried out by hybrid EM algorithm for an efficient computation. We apply our method to real and simulated data set and compare it to two-way mixed ANOVA method. With little model assumptions FPC analysis shows better performances.

RDCurve: A Nonparametric Method To Evaluate the Stability of Selection Procedures

❖ Xin Lu, University of California, San Diego, 9500 Gilman Dr, MC0645, La Jolla, CA 92093, *xinlu@ucsd.edu*

Key Words: gene selection, microarray, re-discovery rate

We proposed a nonparametric Re-Discovery Curve (RDCurve) method, to estimate the probability of re-discovery of gene selection from microarray data sets. Given a selection procedure and a data set, the RDCurve method applies the selection procedure repeatedly to bootstrapped data, select a subset of important genes, and then estimate the expected frequency of rediscovery of the selected subset of genes. We also proposed a permutation method to estimate the confidence band of RDCurve under Null hypothesis to justify the significance of the RDCurve. The method we proposed is a complement to traditional FDR method. It is nonparametric and model independent. With the RDCurve method, we can also estimate the signal-noise ratio of a give data set, compare the performance of two selection procedures in term of them expected rediscovery rate, or select the number of genes to be reported.

Sample Size Estimation for Factorial Designs in Microarray Studies

❖ Bruce Thompson, Clinical Trials & Surveys Corp., 2 Hamill Road, Suite 350, Baltimore, MD 21210, *bthompson@c-tasc.com*; Clara Chu, Clinical Trials & Surveys Corp.; Thomas J. Richards, University of Pittsburgh Medical Center

Key Words: microarray studies, sample size estimation, false positive, false negative

One can use factorial designs for sample size estimation in microarray studies. Sample size calculations follow the multiple testing framework given by Benjamini and Hochberg (1995), which limit the number of false positive and negative determinations of differential expression. The method for a 2X2 design follows: 1) set alpha and beta and calculate a sample size for a single gene analysis such that: among G independent genes (with G0 and G1 genes not differentially and differentially expressed, respectively), the probability that phat (the proportion of genes that test as differentially expressed among the tests showing differential expression) is greater than "p" (e.g., 0.95), will have a specified confidence (e.g., 99%) 2) divide the required sample size in half to obtain the number for each of the four groups. An example will be provided from the Lung Tissue Research Consortium.

Normalization for Replicated MicroRNA Array Data with Spike-In Controls

❖ Yuhua Su, Yale University, 25 York Street, School of Medicine, WWW 218, New Haven, CT 06520, *yuhua.su@yale.edu*; Joanne B. Weidhaas, Yale University; Daniel Zelterman, Yale University

Key Words: microRNA array, normalization, triple-negative breast cancer, spike-in control

Microarray experiments involve multiple readings of gene expression levels and have multiple sources of variation including scanner settings and background readings. Normalization refers to the process of identifying the sources of variation and handling the microarray data in such a way that the expression levels are biologically comparable between arrays or between the printed probes. We describe using the background level readings and spike-in controls to normalize the arrays. This work develops a transformation method to adjust the arrays so that linear relationships among probes will be preserved. The proposed method is illustrated using microRNA array data from a study of triple-negative (HR/PR/ER) breast cancer patients.

144 Machine Learning and Related Issues ● ✪

Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

Integration of Relational and Hierarchical Network Information: Prediction of Protein Function

❖ Xiaoyu Jiang, Boston University, 111 Cummington Street, Dept of Math and Stats Boston University, Boston, MA 02215, *xiaoyu@math.bu.edu*; Eric D. Kolaczyk, Boston University; Simon Kasif, Boston University

Key Words: protein function prediction, Gene Ontology hierarchy, posterior probability, Bayesian learning

In the current climate of high-throughput computational biology, the inference of a protein's function from related covariates, such as protein-protein interaction (PPI) relations, has become a canonical problem. Most existing technologies pursue this task as a Gene Ontology (GO) term-based classification problem. However, ontology structures are essentially hierarchies, with certain top to bottom annotation rules. We propose a probabilistic framework to integrate relational data, in the form of a PPI network, and the GO hierarchy, and offer two classifiers classifying GO terms marginally and jointly. Efficient dynamic programming algorithms for their computation are derived. We apply and evaluate our model in the yeast Saccharomyces cerevisiae for a whole-genome protein function prediction. It is found that substantial improvements may be obtained over non-hierarchical methods.

Subject-Adaptive, Real-Time Sleep Stage Classification Based on Conditional Random Field

❖ Gang Luo, IBM T.J. Watson Research Center, 19 skyline drive, Hawthorne, NY 10532, *luog@us.ibm.com*; Wanli Min, IBM T.J. Watson Research Center

Key Words: pattern recognition, conditional random field, sleep staging, spectrum analysis, subject adaptation, machine learning

Sleep stager performs one of the most important steps in sleep analysis: classifying sleep recordings into sleep stages. We report an online sleep stager using electroencephalogram (EEG) signal based on a recently developed statistical method, conditional random field. Using sleep recordings from human subjects and birds, we show that the average classification accuracy of our sleep stager almost approaches the theoretical limit and is about 8% higher than that of existing systems. Moreover, for a new subject snew with limited training data Dnew, we perform subject adaptation to improve classification accuracy. Using sleep recordings from human subjects, we show that even without any Dnew, our sleep stager can achieve an average classification accuracy of 70% on snew. This accuracy increases with the size of Dnew and eventually becomes close to the theoretical limit.

Variable Selection for Optimal Decisionmaking

❖ Lacey Gunter, University of Michigan, 439 West Hall, Statistics Department, 1085 South University, Ann Arbor, MI 48109, *lgunter@umich.edu*; Susan Murphy, University of Michigan; Ji Zhu, University of Michigan

Key Words: variable selection, decision making, prediction

This talk will discuss variable selection for decision making. Current variable selection techniques were developed for use in a supervised learning setting where the goal is optimal prediction of the response. These techniques often leave behind small but important interaction variables that are critical when the ultimate goal is optimal decision making rather than optimal prediction. While prediction represents a first step in finding optimal decisions, we will point out some key differences between prediction and decision making applications. We will present a new technique designed specifically to find variables that aid in decision making and demonstrate the utility of this technique on both simulated data and real world data from a randomized controlled trial for the treatment of depression.

Correlation Analysis Between Two Sets of Brain Images

❖ Yeming Ma, State University of New York at Stony Brook, 19 Quaker Hill Road, Stony Brook, NY 11790, *jimma@bnl.gov*

Key Words: Correlation, 3D brain image, normalization, PET-FDG, statistical parametric mapping (SPM), voxel

We propose a method in search for brain regions that have significant correlation between two sets of 3D brain images. It extended the capability of the existing SPM software and is generally applicable to any images. Two preprocessing steps were exacerbated from SPM, spatial normalization to a standard brain template and spatial smoothing for multiple comparison correction. We compute the correlation coefficient from the preprocessed images using a "same location strategy" at each voxel to obtain a correlation r-map. Similar to SPM activation concept, activated regions were defined above a threshold calculated from a p-value selected by users. As an example we analyzed brain glucose metabolic images (PET-FDG) taken on treatment and placebo conditions. Correlation maps were computed for 21 cocaine-addictive and 16 healthy control subjects.

Covariate Effect Estimation with Partial Penalization

❖ Jiang Gui, Dartmouth Medical School, One Medical Center Dr, HB 7927, Lebanon, NH 03756, *jiang.gui@dartmouth.edu*; James Sargent, Dartmouth Medical School

Key Words: LARS, LASSO, adolescent alcohol misuse, Movie influence, Penalization, bootstrap

Chronic alcohol use is an important cause of cancers of the digestive tract. Understanding why young adolescents begin to drink and go on to escalate alcohol use is a key element to preventing adolescent alcohol misuse. Although it is known that adolescents initiate drinking in response to social influences, the entertainment industry remains an understudied influence. In this talk, we propose to estimate the effect of movie alcohol exposure while controlling for a number of personal and social risk factors by penalization. We will also estimate the significance of the movie effect via bootstrap.

Variable Importance Selection: Random Forest vs. Logistic Regression

❖ Andrejus Parfionovas, Utah State University, 563 N 700 E apt 9, Logan, UT 84321, *andrej@cc.usu.edu*; Adele Cutler, Utah State University

Key Words: variable selection, Random Forest, Logistic Regression, multicollinearity, classification, cardiac events

We demonstrate the efficiency of variable selection for multivariate analysis using Random Forests (RF) and compare it to Logistic Regression (LR) using numerical simulations. RF demonstrates higher success rate for choosing statistically important variables on highly correlated and/or noisy data. We discovered the essential difference between the variable selection mechanism of two approaches: the RF assigns importance based on its explanatory value of the variable, while LR focuses on a subset that provide sufficient explanatory effect, thus completely ignoring other variables of possible interest. RF provides more informative insight of the data, is more robust and stable. Using Kolmogorov-Smirnov test we propose a comprehensible and easy to use method to compare the variables importance illustrated on a real-life data example (cardiac events prediction).

Inference for Dynamic Treatment Regimes via Q-Learning

❖ Bibhas Chakraborty, University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109-1107, *bibhas@umich.edu*; Susan Murphy, University of Michigan

Key Words: dynamic treatment regime, Q-Learning, longitudinal data, time-varying treatment, chronic disorders

Clinical treatments of many chronic disorders are time-varying and adaptive to an individual patient's changing health status. In this context, a dynamic treatment regime is a list of decision rules that tells how the level of treatment should be tailored through time according to an individual's response to ongoing therapy. In this paper, we consider the problems of estimating dynamic treatment regimes and conducting inference on them from longitudinal data on patients. For estimation, we propose a method called Q-Learning. But this method being non-smooth, neither asymptotic normality nor bootstrap can be used to get valid tests of hypothesis. We illustrate this via simulations. As a remedy, we propose a regularized method and prove that the estimators obtained from this method are asymptotically normal. Performance of this method is assessed through simulations.

145 Statistical Issues in Noninferiority and Equivalence Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Monday, July 30, 10:30 am–12:20 pm

Generalized Mixed Models for Noninferiority Trials

❖ Robert Abugov, Food and Drug Administration, 7500 Standish Place, Rockville, MD 20855, *robert.abugov@fda.hhs.gov*

Key Words: noninferiority, clinical trial, positive control, mixed effect model

Noninferiority trials consider a new therapy effective if the upper confidence bound on the difference in success between positive control and investigational treatment does not exceed a prespecified margin. Such trials commonly employ binary response data collected at multiple sites with multiple investigators per site, suggesting the generalized linear mixed effect model as appropriate for calculating the needed upper confidence bound on difference in success. This Monte Carlo study compares the realized accuracy of different methods for calculating confidence bounds by comparing actual coverage to intended percent confidence. If there is time, we will address model robustness as well as contingency planning should the most appropriate model fail to converge.

Trials and Tribulations in Choosing Noninferiority Margins

❖ John Johnson, Cato Research, Ltd., 4364 South Alston Avenue, Westpark Corporate Center, Durham, NC 27713-2280, *jdjohnso@cato.com*

Key Words: noninferiority trials, margins, clinical trials

The choice of noninferiority margin in clinical trials is under greater scrutiny, but scant guidance is provided from regulatory agencies and the literature on just how to justify the margin. For example, regulatory guidance with specific advice have been withdrawn, and sponsors of trials have been called upon to write more specific justifications in their protocols. Some difficulties encountered during a recent case are presented, along with a discussion of how they were handled.

A Testing Strategy for Noninferiority and/or Superiority Hypotheses on Multiple Endpoints

❖ Xiaohui Luo, Merck & Co., Inc., 181 Buttonwood Dr, Piscataway, NJ 08854, *edmund_luo@merck.com*; Jianjun Li, Merck & Co., Inc.

Key Words: Multiple comparison, Type I error, Power, Dose Range Finding, Active control, Non-inferior

In clinical trials comparing a study medication with an active control on multiple endpoints, it is not always the case that we need to show superiority of the study medication to the active control on all tested endpoints. Instead, the trial can be claimed positive if we can show that compared with the active control, the study medication is non-inferior on all endpoints, and superior on at least one of the endpoints. For example, the study medication is non-inferior on one efficacy endpoint and superior on the safety endpoint. We propose several approaches to test this composite hypothesis on multiple endpoints. Based on the comparison of our proposed approaches and the methods in the literature, the optimal testing strategy will be recommended and the formula for power calculation will be developed. The strategy will also be developed for dose range finding studies.

New Tests for Assessing Noninferiority and Equivalence from Survival Data

❖ Kallappa Koti, Food and Drug Administration, DBV, 10903 New Hampshire Avenue, Silver Spring, MD 20993, *kallappa.koti@fda.hhs.gov*

Key Words: Two-arm trials, Right-censored data, Kaplan-Meier estimate, Generic drugs

We propose a new method for assessing non-inferiority of an experimental therapy compared to a standard of care. The ratio μE/μR of true median survival times is the parameter of interest. We think of the ratio mE/mR of the sample medians as a point estimate of the ratio μE/μR. We use the Fieller-Hinkley distribution of the ratio of two normally distributed random variables to derive a level-a test of inferiority null hypothesis, which is stated in terms of the ratio μE/μR and a pre-specified fixed non-inferiority margin. We also explain how to assess equivalence and non-inferiority using bootstrap equivalent confidence intervals on the ratio μE/μR. We discuss sample size determination. The proposed new test does not require the hazard rates to be proportional. We claim that our test procedure is simple and attains adequate power for moderate sample sizes.

A Multivariate Test for Population Bioequivalence

❖Inna Chervoneva, Thomas Jefferson University, 1015 Chestnut St Suite M100, Philadelphia, PA 19107, *i_chervoneva@mail.jci.tju.edu*; Terry Hyslop, Thomas Jefferson University; Walter W. Hauck, Kimmel Cancer Center at Jefferson

Key Words: Multivariate equivalence, Population bioequivalence, In vitro bioequivalence, Modified large-sample confidence intervals

We propose a multivariate generalization of the univariate criteria for testing population bioequivalence. Most approaches for testing multivariate equivalence consider a multivariate equivalence region, which implies simultaneous comparison of means in each dimension. In contrast, our proposal combines a comparison of means and a comparison of variances into a single aggregate criterion, using the trace of the covariance matrix as a scalar measure of the total variability. We use a confidence interval approach to testing and construct two versions of the modified large-sample confidence interval for the linearized multivariate criterion. In a simulation study, we evaluate the empirical coverage of these confidence intervals and rejection rates of the corresponding tests. The proposed methodology is illustrated with an example of testing equivalence of the spray pattern of nasal sprays.

Estimation of Long-Term Vaccine Efficacy and Waning Effects Using Surveillance Interval Data

❖Martha Lee, Food and Drug Administration, 1401 Rockville Pike, HFM217 WOC1 FDA CBER, Rockville, MD 20852, *martha.lee@fda.hhs.gov*; Tsai-Lien Lin, Food and Drug Administration

Key Words: vaccine efficacy, waning effects, surveillance, interval data, breakthrough case

Introduction: In light of logistical and financial limitations of clinical trials for long periods of follow-up, epidemiologic data are often used in estimating long-term vaccine efficacy and waning effects. Many mathematical models have been proposed to investigate waning vaccine efficacy. In this study, we focus on the analytic methods for interval data collected through surveillance system. Methods: Breakthrough cases can be identified in surveillance system. We adapt the basic concept of non-inferiority to test effect preservation between two adjunct surveillance intervals. Waning effects can be detected as effect preservation fails. Results: We apply the proposed methods to both real-world and simulated datasets. Conclusion: This study is of importance in evaluating long-term vaccine efficacy.

An Equivalence Test Based on the Duration of Treatment Effect as Measured by Median Time-to-Event

❖Suriani Pokta, Allergan, Inc., 2525 Dupont Drive, MS MIC1H (Mail code: MI2-1M), Irvine, CA 92612, *pokta_suriani@allergan.com*; Jihao Zhou, Allergan, Inc.

Key Words: equivalence test, median time-to-event, duration of treatment effect

In clinical trials involving systemic compounds, bioequivalence (BE) studies are usually adequate to demonstrate the therapeutic equivalence between a new formulation and a standard formulation, or between a generic drug and a reference-listed drug. However, BE studies are irrelevant for topical compounds. Therefore, clinical equivalence trials are required. In this paper we present a new equivalence test based on the duration of treatment effect as measured by median time-to-event. First, we derive a parametric equivalence test based on an exponential survival distribution with censoring. Then we develop a nonparametric equivalence test based on median Kaplan-Meier estimates of the duration of treatment effect. The performance of these tests is assessed through simulated data. For illustration, we apply the above methods to a published clinical trial.

146 Long Memory and Related Processes

Business and Economics Statistics Section

Monday, July 30, 10:30 am–12:20 pm

Estimation in Long-Time Memory Processes

❖Wen Li, Iowa State University, Apt 9, 230 Raphael Ave, Ames, IA 50010, *shirley@iastate.edu*; Alicia Carriquiry, Iowa State University; Wolfgang Kliemann, Iowa State University; Cindy Yu, Iowa State University

Key Words: ARFIMA model, Autocorrelation, Hurst exponent, Self-similar index, Long-time memory

When the correlation between current and lagged observations decays very slowly, data collected over time often have the property of long memory, which has been characterized in various ways in the literature. Perhaps the four most popular characterizations of long memory are based on the autocorrelation function, on the self-similarity index, on ARFIMA-type models and on the Hurst exponent. In this paper, we investigate relationships between those four approaches and derive some of their properties. Further, we focus on several estimators of long memory and via simulation estimate the bias, sampling standard deviation and root mean squared error under a set of true values of the Hurst exponent. Finally, we compare the performance of several estimators of Hurst exponent by applying them to a subset of the S&P 500 series and without relying on an underlying dynamic model.

Unit Root Tests in the Presence of a Simultaneous Break in the Mean and the Innovation Variance

❖Amit Sen, Xavier University, Department of Economics, 3800 Victory Parkway, Cincinnati, OH 45207-3212, *sen@xavier.edu*

Key Words: Unit Root, Break Date, Innovation Break

Kim, Leybourne, and Newbold (2002, Journal of Econometrics) developed unit root tests that allow for a break in the innovation variance. Their tests are based on a consistent break-date estimator, and a modified GLS regression using the pre-break and post-break variance implied by the estimated break-date. We show that the break-date estimator of Kim, Leybourne, and Newbold (2002) is not consistent if there is a break in the mean under the alternative, and so we propose a modified estimator for the break-date. Implementing Kim, Leybourne, and Newbold's (2002) GLS strategy with the modified break-date estimator yields appropriate unit root tests in the presence of a break in the innovation variance that will also have power against the stationary alternative with a break in the mean. Application of the unit-root test is illustrated via application to the Thai Baht/U.S. $ exchange rate.

Long-Memory Parameter Estimation in Time Series and Its Connection to fBm

❖Michael Levine, Purdue University, Dept of Statistics, 250 N University St, West Lafayette, IN 47907, *mlevins@stat.purdue.edu*; Frederi Viens, Purdue University; Soledad Torres, Universidad de Valparaiso

Key Words: Hurst parameter, LARCH, Conditional MLE, Long-memory property

We investigate several possible strategies for consistently estimating the Hurst parameter H responsible for the long-memory property in a special class of nonlinear time series ARCH-type models known as LARCH, as well as in the continuous-time fractional Brownian motion (fBm). Conditional MLE estimation method and a local Whittle-type estimation procedure for this parameter are discussed. The conditional MLE is proved to be consistent and a Portmanteau-type test for model validation is established. A specially designed conditional maximum likelihood method for estimating the fBm's Hurst parameter is proposed. In keeping with the popular financial interpretation of ARCH-type models, all estimators are based only on observation of the returns of the model and not on the volatilities.

A Power Study of Seasonal Unit Root Tests for Quarterly Data

❖Qianyi Zhang, North Carolina State University, 3520 Cum Laude Ct, Apartment 301, Raleigh, NC 27606, *qzhang3@ncsu.edu*; David Dickey, North Carolina State University; Sastry Pantula, North Carolina State University

Key Words: seasonal unit root, frequency, power

There are two real unit roots (1 and -1) and a pair of conjugate unit roots (i and -i) for quarterly processes with seasonal unit roots at all frequencies. This paper reveals why the test of Dickey, Hasza and Fuller (DHF), which tests the null hypothesis that the characteristic equation has all 4 of these roots, has low power when the underlying processes have seasonal unit roots at some but not all of these frequencies. By focusing on the time series with seasonal unit roots only at zero and semi-annual frequency (1 and -1), this study proposes a test of the null hypothesis of seasonal unit roots at all frequencies against the alternative of seasonal unit roots of 1 and -1 only. Monte Carlo simulations are used to compare the performance of DHF, Kunst test and the proposed test.

Test of Cointegration Using Long Run Canonical Correlations: Part I

❖Kalidas Jana, University of Texas at Brownsville, Dept of Bus Admin, 80 Fort Brown, Brownsville, TX 78520, *kalidas.jana@utb.edu*; Alastair R. Hall, University of Manchester

Key Words: Long Run Canonical Correlations, Cointegration, Squared Coherence at Frequency Zero, Squared Multiple Coherence at Frequency Zero, Squared Canonical Coherences at Frequency Zero

Recently, Long Run Canonical Correlations (LRCCs) have been shown to be useful in providing a suitable metric for information contained in the population moment conditions in Generalized Method of Moments (GMM) estimation. LRCCs have also been shown to be useful in structural stability testing, and exogeneity testing of regressors when the regressors are non-stationary. This paper explores further usefulness of LRCCs. In particular, it investigates the implications of LRCCs for cointegration. In addition, exploiting these implications, it develops new tests of cointegration in the frequency domain analysis of time series. Depending on the situation at hand, the tests involve squared coherence at frequency zero, squared multiple coherence at frequency zero, or squared canonical coherences at frequency zero.

Contemporaneous Aggregation of Time Series

❖Georgios Tripodis, University of Massachusetts Amherst, 1040 N Pleasant Street, Apt 324, Amherst, MA 01002, *tripodis@math.umass.edu*; Jeremy Penzer, London School of Economics

Key Words: aggregation, unobserved component models, Kalman filter

This paper looks at the problem of contemporaneous aggregation of time series. Extensive work has been done in this area for ARIMA models. We consider the unobserved components framework. We consider different mechanisms for aggregating time series models and derive the conditions for identifiability for the aggregate series. We show that identifiability of the models for the component series is not sufficient for the identifiability of the model for the aggregate series. We also consider the case where there is no estimation error as well as the case of modeling an unknown process. For the case of the unknown process we provide recursions based on the Kalman filter that give the asymptotic variance of the estimated parameters.

Feature Extraction Using Functional Data Analysis for Screening Large Numbers of Time Series

❖Ganesh Subramaniam, AT&T Labs - Research, 180 Park avenue, Room A251 Building 103, Florham Park, NJ 07932, *mkg@research.att.com*; Ravi Varadhan, Johns Hopkins University; Ciprian M. Crainiceanu, Johns Hopkins University

Key Words: Semiparametric Models, Smoothing, Penalized Splines, Derivatives, Data Mining

Business decisions and business process monitoring are often based on time series data that represent an aggregation of a large number of time series. Although the inferences are generally based on the aggregate data, significant insights lurk in the underlying time series that had been combined. The challenge that analysts currently face is the large number and complexity of data underlying aggregate time series. This paper provides a framework that supports drill down analysis and screening of large scale time series data by developing feature extraction rules. We develop an exploratory method based on functional data analysis, where we fit smooth functions with semiparametric models using penalized splines. One feature extraction involves estimating derivatives from these models. These provide insights on the bumps and dips for the underlying time series.

147 Bayesian Approaches in Machine Learning and Model Selection

Section on Bayesian Statistical Science

Monday, July 30, 10:30 am–12:20 pm

Principal Component Reduction in Linear Mixed B-Splines

❖Carsten Botts, Williams College, 96 School St 10, Williamstown, MA 01267, *cbotts@williams.edu*; Michael Daniels, University of Florida

Key Words: Laplace Approximation, Unit-Information Prior, Reversible jump MCMC

We consider a linear mixed effects model for longitudinal trajectories. In this mixed model, the population curve and the subject specific deviates are both modeled as a b-spline with k knots located at t_k. We propose a novel Bayesian method of identifying possible values of k and t_k while also reducing the number of principal component curves associated with the random effects. Specifically, we use reversible jump MCMC methods to sample from a posterior distribution of k, t_k, and r, where r is the number of principal component curves retained. This method of knot selection and principal component reduction is studied through simulations and then applied to a dataset.

Adaptive Dynamic Bayesian Networks

❖Brenda Ng, Lawrence Livermore National Laboratory, P.O. Box 808, L-229, Livermore, CA 94551, *bmng@llnl.gov*

Key Words: adaptive models, dynamic Bayesian networks, nonparametric Bayesian modeling, graphical models

A discrete-time Markov process can be compactly modeled as a dynamic Bayesian network (DBN)–a graphical model with nodes representing random variables and directed arcs indicating causality between variables. Each node has a probability distribution, conditional on the variables represented by the parent nodes. A DBN's graphical structure encodes fixed conditional dependencies between variables. But in real-world systems, conditional dependencies between variables may vary over time. Model errors can result if the DBN fails to capture all possible interactions between variables. Thus we propose adaptive DBNs, whose structure and parameters can change: a distribution's parameters and its set of conditional variables are dynamic. This work builds on nonparametric Bayesian modeling and machine learning methods, such as structural EM. We show its advantages in a complex plant modeling task.

Testing Equality of Several Functions

❖Sam Behseta, California State University, 9001 Stockdale Hwy, Mathematics Department, Bakersfield, CA 93311, *sbehseta@csub.edu*; Robert E. Kass, Carnegie Mellon University

Key Words: Bayesian Functional Data Analysis, Bayesian Adaptive Regression Splines, Likelihood Ratio Tests, Analysis of Neuronal Data

We consider the problem of comparing several noisy functions. We propose testing procedures to perform comparisons either pointwise or globally over the entire length of functions. A simulation study of power demonstrates the strength of these tests even for moderate sample sizes. We implement these tests on a group of neurons recorded from primate frontal oculomotor cortex, first to screen for condition-related differential activity and, second, to search for neurons displaying interesting patterns that vary with condition.

Aspects of Bayesian Lasso Regression

❖Christopher Hans, The Ohio State University, 1958 Neil Avenue, Cockins Hall Room 404, Columbus, OH 43210-1247, *hans@stat.osu.edu*; Steven N. MacEachern, The Ohio State University

Key Words: Bayesian regression, Gibbs sampler, lasso, MCMC, prediction

Connections between Lasso estimates for linear regression and Bayesian methods usually focus on the posterior mode even though the Bayes estimate under squared error loss – perhaps the most commonly used loss function for Bayesian estimation – is the posterior mean. We present aspects of Bayesian Lasso regression that focus on parameter estimation via the posterior mean and on prediction of future cases via the posterior predictive distribution. Estimation of the full posterior distribution is accomplished through component-wise Gibbs sampling, including transform methods to improve mixing. While inference based on the posterior mean does not perform "variable selection" in the same way as the usual Lasso estimate, we address the question of Bayesian model comparison and selection. A link to software with efficient implementation of the methods with a simple R interface is provided.

Bayesian Adaptive LASSO for Linear Models

❖Yi-Liang Tung, National Cheng Kung University, No1 University Road, Tainan City, 701 Taiwan, *tung.yiliang@gmail.com*

Key Words: Model averaging, Model selection, Penalized least squares, Shooting algorithm, Shrinkage

We consider the problem of selection of informative covariates in linear regression models. We introduce a novel method to select variables, which is referred to as the Bayesian adaptive LASSO. This proposal is motivated by a particular hierarchical Bayesian model that is able to provide adaptive information to identify important covariates to be included in the final model. Then, we adopt the posterior information to construct an estimation criterion that shares the key feature of the adaptive LASSO. In particular, we design an efficient MCMC algorithm to handle data sets involving a large number of covariates. We compare the proposal with its several competitors via simulation experiments. The results show that the proposal performs better in terms of the prediction accuracy and the relative model error. Finally, the proposed method is illustrated with two real-world datasets.

Bayesian Kernel Regression with Feature Selection

❖Zhi Ouyang, Duke University, 4225 Larchmont Rd, Apt 537, Durham, NC 27707, *zo2@stat.duke.edu*; Merlise A. Clyde, Duke University; Robert L. Wolpert, Duke University

Key Words: Bayesian kernel regression, feature selection, sparsity, Lévy processes, support vector machine, relevance vector machine

We propose a general Bayesian framework to make feature selection using kernel regression. Unlike most kernel regression models in literature, our model allow kernel functions to have different shape parameters. Those shape parameters give us great flexibility to model different types of data. The approach uses Lévy priors to promote sparsity in both the kernel functions and predictor variables (or features); these priors act as regularizers for the likelihood function that reward good selected features. In large p small n problems, we keep the kernel centers within the observed data points, however, we could have more than one kernel centered at the same location. We relate this method to other work such as Support Vector Machines (SVM) and Relevance Vector Machine (RVM).

Bayesian Model Averaging Using Adaptive Sampling

❖Joyee Ghosh, Duke University, Box 90251, Durham, NC 27708-0251, *joyee@stat.duke.edu*; Merlise A. Clyde, Duke University

Key Words: Bayesian Model Averaging, Sampling Without Replacement, Two Urn Model

Implementing Bayesian Model Averaging (BMA) when the number of explanatory variables, p, is large is a challenging task. Our focus will be on problems where the marginal likelihoods can be calculated analytically. When p is greater than 25 calculating the marginal likelihoods of all models becomes computationally intractable and typically BMA is based on a subset of models. We use an adaptive sampling algorithm due to Clyde and Littman to choose this subset which samples models without replacement. Our motivation for sampling without replacement is that once we sample a model we can calculate its marginal likelihood analytically and would not like to revisit it. We use some ideas from the Bayesian approach to finite population sampling to estimate the posterior probability of the unsampled models, based on this we decide whether we need to continue sampling.

148 Innovative Bayesian and Frequentist Models for Health Using Individual and Ecologic Data ● ✪

Section on Health Policy Statistics, ENAR, Biometrics Section, WNAR

Monday, July 30, 10:30 am–12:20 pm

Truth or Consequences: Estimating Error Due to Inconsistent Self-Report of Adolescent Risk Behaviors

❖Janet Rosenbaum, Harvard University, 15 Leonard Av, Cambridge, MA 02139, *janet@post.harvard.edu*

Key Words: adolescents, risk behavior, reliability, Bayesian, test-retest, YRBS

This study identifies question properties associated with giving logically inconsistent answers to survey questions using a test-retest study of the Youth Risk Behavior Survey (YRBS). High school students answered the same 72 questions at an interval of two weeks (n=4628). The prevalence of 41 of the 72 risk behaviors changed significantly in 2 weeks. Questions about substance use are more consistent than other topics. Questions about lifetime risk behaviors are more consistent than questions about behaviors in other time frames. A Bayesian method estimated that standard error of prevalence of 72 risk behaviors is at median 3.3 times larger than conventional estimates of standard error (95% CI (1.9, 5.0)). Most questions on YRBS do not yield reliable estimates of risk behavior prevalence. Ambiguous questions are more inconsistent than questions with a single interpretation.

A Statistical Tool Based on a Model of Personal and Ecologic Characteristics To Explain Deprivation in U.S. Health Care Access

❖Martey Dodoo, The Robert Graham Center, 1350 Connecticut Ave NW, Suite 201, Washington, DC 20036, *mdodoo@aafp.org*; Robert Phillips, Robert Graham Center; Xingyou Zhang, The Robert Graham Center

Key Words: health, access, deprivation

There have been attempts using personal characteristics to explain barriers/delays in accessing healthcare, and using ecological measures as near-proxies for health access problems. We combined these efforts to develop indices of health access deprivation using 2001-2003 NHIS data, and constructing indices using PCA. We merged indices with census tract level data, and used multi-level logistic modeling to pair wise substitute Census for NHIS variables. We found 10 predictors of health deprivation out of 23 potential from literature. Deprivation is more likely, if one is female, do not own a home, live in a single adult household, have a functional limitation, smoke, have no health insurance or is less than 65. We mapped indices and overlaid them with provider availability statistics. This tool could assist local community assessments, health planning, and identifying populations at risk.

Bayesian Hierarchal Modeling of Small-Area Suicide Rates: A Geographical and Ecological Evaluation

❖Yi Lin, University of British Columbia, Room E414, 4480 Oak Street, Vancouver, BC V6H3V4 Canada, *yxlin@cw.bc.ca*; Ying MacNab, University of British Columbia

Key Words: Bayesian hierarchical model, shared component model, conditional autoregressive model, suicide rates, geographical variation, ecological analysis

We present a Bayesian spatial and ecological analysis of hospitalization and mortality rates of suicide among children and youth in British Columbia, Canada. The Bayesian spatial modeling enabled us to produce more reliable risk estimates and to identify regional level risk factors. Spatial and shared component models were used to explore and quantify spatial correlation in relative risk and shared component effects. Empirical Bayes penalized quasi-likelihood method is used to identify risk factors and fully Bayesian MCMC method is used to quantify uncertainties associated with the estimation of all model parameters, particular the regional relative risks. Principal component method was used to explore multicollinearity among covariates and to control for confounding. Regional characteristics explained considerable variations in suicide rates.

Distribution of Cronbach's Alpha for Ordinal Data: A Bayesian-Based Approach

❖Byron Gajewski, The University of Kansas Medical Center, 3901 Rainbow Blvd, MS 4043, Kansas City, KS 66160, *bgajewski@kumc.edu*; Diane K. Boyle, The University of Kansas Medical Center; Sarah Thompson, University of Nebraska Medical Center

Key Words: Confidence intervals, Credible intervals, latent variables, psychometric evaluation, item response models, multivariate normal distribution

We propose point estimates and intervals for Cronbach's alpha when the data are ordinal. Traditional calculations of Cronbach's alpha on ordinal instruments underestimate the true Cronbach's alpha provided from the latent variables that are assumed to produce the ordinal data. By utilizing Bayesian models and multiple imputation, confidence (credible) intervals are provided for Cronbach's alpha at the latent variable level. The proposed methodology is shown to have theoretically correct coverage probability and is demonstrated on an instrument that measures nursing home residents' quality of life.

A Bayesian Two-Part Model for Bounded Non-Negative Data: Estimating Extra Time Spent on Diabetes Self-Care

❖Betsy C. Gunnels, Centers for Disease Control and Prevention, 3470 Buford Highway MS K10, Atlanta, GA 30341, *bic6@cdc.gov*; Theodore J. Thompson, Centers for Disease Control and Prevention; Louise B. Russell, Rutgers University; Susan L. Ettner, University of California, Los Angeles; James P. Boyle, Centers for Disease Control and Prevention

Key Words: Two-part models, Bayesian modeling, Diabetes Mellitus

Non-negative data with a large proportion of zeros is a candidate for two-part modeling. Two-part models consist of a binary regression distinguishing zeros from positive values and a linear regression for the (transformed) positive values. We extend this approach in two ways. We use a t-distribution for the errors in the linear regression instead of the normal distribution. The t-distribution is an option for modeling heteroscedastic normal data. We also consider a set of transformations that constrain the retransformed predictions to be bounded. Candidates include versions of the logistic, probit, complementary log-log, and log-log. Cross validation of posterior predictions is used to compare models and evaluate the fit of the final model. We demonstrate the methods using extra time spent per day on diabetes self care as our outcome with income and education as main exposures.

Joint Modeling of Birth Weight and Gestational Age via the Gibbs Sampler

❖Betsy Enstrom, Duke University, 2555 NE Loop 410 Apt 706, San Antonio, TX 78217, *betsy@stat.duke.edu*; Alan E. Gelfand, Duke University; Geeta Swamy, Duke University; Marie Lynn Miranda, Duke University

Key Words: mixture model, label-switching, Gibbs Sampler, MCMC algorithm

Modeling birth weight and gestational age have received much discussion in the literature. Linkage to a wide range of social, environmental, and genetic risk factors is of primary interest. We propose joint modeling with the objective of better characterization of an adverse birth outcome. We also note that, in building of a joint model, we will gain the understanding of both of the cohorts and their correlation. Applying these models to a data set consisting 107045 live births as recorded in the North Carolina Detailed Birth Record for 2003, we develop the model specifications in a hierarchical Bayesian framework. A Gibbs sampler is illustrated for the Bayesian approach with a mixture model of bivariate Gaussians. Practical issues that arise when fitting the multivariate mixture model of birth weight and gestational age with a Bayesian approach are discussed here.

Characterizing Drug Diffusion with Patient-Level Data

❖Camelia Sima, MSKCC, 307 E 63rd Street 3rd floor, New York, NY 10021, *simac@mskcc.org*; Katherine Panageas, MSKCC; Glenn Heller, MSKCC; Deborah Schrag, MSKCC

Key Words: drug diffusion, competing risks, bivariate cumulative incidence function

To inform assessments of the quality of cancer care, we investigate different trends of chemotherapy drugs subsequent to FDA approval, as functions of tumor type and patient characteristics. The event of interest is the time to first drug utilization by an eligible patient, and the analysis is complicated by dependent censoring for death. Diffusion is estimated through a bivariate cumulative incidence function, accounting for death as a competing risk. The two components are time from cancer diagnosis to first utilization of the drug, and calendar time. The relation between the components is of interest in order to understand how drug adoption is related to the time elapsed since FDA approval. We apply the method to SEER-Medicare data and test the hypothesis that worse prognosis and younger age at diagnosis are correlated with accelerated drug diffusion.

149 Semiparametric and Nonparametric Inference ✪

Section on Nonparametric Statistics

Monday, July 30, 10:30 am–12:20 pm

Confidence Intervals for a Discrete Population Median

❖Ronald Randles, University of Florida, PO Box 118545, Gainesville, FL 32611-8545, *rrandles@stat.ufl.edu*; Denis Larocque, HEC Montreal

Key Words: Median, Confidence Intervals, Discrete

Confidence intervals for a discrete population median which use order statistics as endpoints are described. They differ in the level of confidence attributed to the interval. Their performances are compared via simulation.

A Family of Kurtosis Orderings for Multivariate Distributions

❖Jin Wang, Northern Arizona University, Department of Mathematics and Statistics, Flagstaff, AZ 86011-5717, *jin.wang@nau.edu*

Key Words: Kurtosis, Peakedness, Tailweight, Ordering, Elliptically symmetric distributions, Multivariate analysis

Extending the van Zwet (1964) kurtosis ordering for univariate symmetric distributions, we define and study a family of kurtosis orderings for multivariate distributions. All those orderings are affine invariant and that a distribution F is less than or equal to a distribution G in each ordering implies that G has at least as much peakedness and at least as much tailweight as F. All even moments of the Mahalanobis distance of a random vector (if exist) preserve some set of the orderings. For elliptically symmetric distributions, each ordering determines the distribution up to affine equivalence. Ordering results are established for three important families of elliptically symmetric distributions: Kotz type distributions, Pearson Type VII distributions, and Pearson Type II distributions. Finally application to assess multivariate normality is discussed.

Rank-Weighted, Two-Sample U-Statistics with Application to Partial Area Under ROC Curves

❖Chengqing (Alan) Wu, National Institutes of Health, NICHD, 6100 Executive Blvd Rm 7b05, Rockville, MD 20852, *wuch@mail.nih.gov*; Aiyi Liu, National Institutes of Health; Kai Fun Yu, National Institutes of Health

Key Words: Rank Statistics, U-Statistics, Roc curve, Area under ROC curves

A rank weighted two sample U-statistic based on two independent random samples $X_1,\cdots,X_m$ and $Y_1,\cdots, Y_n$ has the form $U_{m,n}=\frac1{mn}\sum_{i=1}^m \sum_{j=1}^n K(X_i,Y_j)J(\frac{R_{xi}}m,\frac{R_{yj}}n) $ where K is fixed measurable function ,$J:[0,1]^2 \rightarrow {\mathcal R} $ is a bounded measurable function and R_{Xi} is the rank order of X_i among X's, R_{yj} is the rank order of Y_j among Y's. A large class of statistics can be expressed as this kind U-statistics or variations thereof. This paper investigates the asymptotic properties of $U_{m,n}$.

When Does the Wilcoxon-Mann-Whitney Fail?

❖Abu Minhajuddin, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Mail Code 8822, Dallas, TX 75390, *abu.minhajuddin@utsouthwestern.edu*; Paul Nakonezny, The University of Texas Southwestern Medical Center at Dallas

Key Words: nonparametric methods, Wilcoxon-Mann-Whitney, Student's t test, two-sample tests, robustness

Wilcoxon-Mann-Whitney (WMW) is perhaps one of the most widely used tests in applied statistics. It is well-accepted as the nonparametric alternative to the Student's two-sample t-test. However, under certain conditions, the WMW may not be the most appropriate test to use. Thus, we explore the general question: "when does the WMW test fail?" Specifically, we address: (1) what are some of the conditions under which the WMW may not be the optimum test of choice? and (2) What is the alternative test in such situations?

Smoothed Mann-Whitney Method for Two-Sample Location Problem

❖Feridun Tasdan, Western Illinois University, Department of Mathematics, Macomb, IL 61455, *f-tasdan@wiu.edu*

Key Words: Shift Parameter, Robust Estimation, Two Sample Problem, Location Parameter

This study is mainly concerned about estimating a shift parameter in the two-sample location problem. Anderson and Hettmansperger (1996) proposed a new method called Generalized Wilcoxon which uses a nondecreasing, and odd function ψ for the pairwise differences of residuals. The proposed Smoothed Mann-Whitney method uses convolution technique for smoothing the empirical distribution functions of the each sample. Then it replaces the empirical cumulative distribution functions with these new smoothed distribution functions. The unknown shift parameter is estimated by solving gradient function with respect to an arbitrary shift variable. Asymptotic properties of the new estimator are established under the conditions that are similar to those used in Generalized Wilcoxon.

Diagnostics for Rank Estimates for a Mixed Model

❖John Kloke, Pomona College, Claremont, CA 91711, *john.kloke@pomona.edu*; Joseph W. McKean, Western Michigan University; M. Mushfiqur Rashid, Food and Drug Administration

Jurevckova (1971) and Jaeckel (1972) developed asymptotic theory for iid linear models. Ordinary rank (OR) estimators are those which minimize Jaeckel's dispersion function. Theory for OR estimators for certain mixed models follows in a manner similar to that of the iid case. We briefly discuss these asymptotic results. The majority of the talk, however, will be devoted to discussing diagnostic procedures for these models (e.g., studentized residuals). References: Jaeckel, L. A. (1972), Estimating regression coefficients by minimizing the dispersion of the residuals, Annals of Mathematical Statistics, 43, 1449–1458. Jurevckova, J. (1971), Nonparametric estimate of regression coefficients, Annals of Mathematical Statistics, 42, 1328–1338.

Clusterwise Regression Using Dirichlet Mixtures

❖Subhashis Ghoshal, North Carolina State University, 2501 Founders Drive, 12 Patterson Hall, Raleigh, NC 27695-8203, *ghoshal@stat.ncsu.edu*; Changku Kang, Bank of Korea

Key Words: Nonparametric regression, clustering, Dirichlet mixture process, MCMC, curse of dimensionality

We consider a clustering based approach to non-parametric regression when data come from finitely many hidden sub-populations in each of which a simple parametric regression model holds. We try to recover the lost labels by a Bayesian clustering technique based on Dirichlet mixture of normals. Because sample is not split regionwise, we avoid the curse of dimensionality problem in higher dimension. The clusters are formed automatically within an MCMC scheme. Model parameters are estimated by least square method in each cluster. An ensemble of parametric regression estimates are formed, each based on a configuration formed in each MCMC step, and a simple averaging produces the final estimate. Our method also gives confidence bands and compares favorably with kernel, spline or GAM based method when clusters overlap less. The method is applied to analyze a donation data.

150 Symbolic, Time Series, and Image Analysis II ● ✪

Section on Statistical Computing, Section on Statistical Graphics
Monday, July 30, 10:30 am–12:20 pm

Symbolic Data Analysis

❖Lynne Billard, University of Georgia, Department of Statistics, Athens, GA 30602-1952, *lynne@stat.uga.edu*

Key Words: symbolic-data, intervals, distributions, variations, classical comparisions

With the advent of the modern computer, contemporary datasets can be too large to handle or analyze in traditional ways. An alternative approach is to summarize the data in such a way that ensures the dataset is more manageable yet also retains as much of the knowledge in the original dataset as possible. One consequence is that the dataset no longer consists of single point values in p-dimensional space but may be lists, intervals, distributions and so on, and are now hypercubes in space. Such data are called symbolic data. Some data are naturally symbolic in nature. We consider how such data arise, how they are structured, and how they can be analyzed. A comparison of symbolic analyses with standard analyses shows that the symbolic methods capture more of the inherent information contained within the symbolic data.

Temporal Statistics for Consequences of Alcohol Use

❖Peter Mburu, George Mason University, 4215 University Drive, Fairfax, VA 22030, *pmburu@gmu.edu*; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University

Key Words: Acute Outcomes, Intervention Strategies, DWI, Time Series

Time-of-day, day-of-week, and month-of-year statistics are important for modeling the acute consequences of alcohol use and abuse. These consequences are not homogeneous through the year, although that assumption is often made when planning intervention strategies. In this paper we consider data from the Virginia Department of Motor Vehicles concerning alcohol-related fatal crashes, also Fairfax County, Virginia DWI arrests and finally Fairfax County, Virginia alcohol treatment admission data. We investigate the cyclic effects over the period 2000–2005.

Time-Frequency Analysis of Electroencephalogram Series

❖Wei Yang, SUNY at Albany, 12144, *wy4515@albany.edu*; Stephen Wong, University of Pennsylvania; Igor Zurbenko, SUNY at Albany

Key Words: EEG, Seizure, Epilepsy, KZFT, Time-Frequency analysis, PPF

The Kolmogorov-Zurbenko Fourier transform (KZFT) is an iterated Fourier transform which may overcome the spectral leakage of regular Fourier transform. Simulation studies have shown its superior resolution over the regular Fourier transform. In this paper, KZFT based time-frequency analysis was applied to EEG series from scalp recordings of generalized tonic-clonic seizures. The time-frequency plot showed that the peak power frequency (PPF) slowly decreased within the theta band before the seizure onset and quickly decreased within the delta band during and after the seizure onset. The estimated decreasing rates were 0.014 and 0.09 Hz/second, respectively. In addition, we observed a trend of increasing PPF of delta band activity prior to seizure onset. These findings may contribute to the identification of seizure onset.

Temporal Extensions to Spatial Statistical Metrics

❖James Shine, U.S. Army Topographic Engineering Center, 4218 Alcott Street, Alexandria, VA 22309, *jshine@tec.army.mil*; James P. Rogers, U.S. Army Corps of Engineers; Mete Celik, The University of Minnesota; Shashi Shekhar, The University of Minnesota

Key Words: spatial statistics, spatio-temporal, point patterns, lattices, random fields

A wealth of metrics have been developed to measure the statistical properties of different types of spatial data. Point patterns are described by metrics such as Ripley's K, intensity, and different nearest-neighbor measures. Lattices are described by metrics such as Moran's I, Geary's C and different forms of spatial regression models. Continuous random field models are described by metrics such as the variogram. As spatial research increasingly becomes spatio-temporal research, extensions of these spatial statistics to the temporal realm are being examined. In particular, extensions to Ripley's K have been discussed by Diggle and others. This paper will examine existing efforts to extend these spatial metrics to include the temporal dimension and will offer some new extension ideas as well.

Approaches to Time Series Clustering

❖Hwanseok Choi, The University of Alabama, PO Box 870221, Tuscaloosa, AL 35487-0221, *hchoi@cba.ua.edu*; J. Michael Hardin, The University of Alabama

Key Words: Time Series Clustering, MPCA, Multivariate Time series, Dimension Extraction

Time series clustering has been one of the major techniques in providing useful information in various fields these days especially as applied to data mining. There seems to be an increased interest in time series clustering in various domains. However, the majority of time series clustering studies are restricted to univariate time series; and, there is lack of studies to compare different time series clustering approaches. In this presentation we introduce several feature extraction approach models to reduce original time series dimensions. These techniques can be used to draw lower dimension clustering plots with multivariate time series data. After that, we compare the results of predictability among multivariate time series clustering methods with simulation data as well as field data. Further, we examine these cluster methodologies across different time series contexts.

An Empirical Spectral Test (EST) for Random Sequences

❖David Zeitler, Grand Valley State University, 1 Campus Drive, 1172 Mackinac, Allendale, MI 49401, *zeitlerd@gvsu.edu*; Joseph W. McKean, Western Michigan University; John Kapenga, Western Michigan University

Key Words: empirical spectral test, random number generator, random sequence test, Fourier transform

Random number sequences are critical to research in many fields. The validity of research results often depend heavily on the underlying distribution of these sequences. In this work we introduce the Empirical Spectral Test (EST). The EST is a highly flexible test of spatial uniformity based on a multi-dimensional Fourier transform of the empirical probability density function. It has properties in common with theoretical tests such as the spectral test and discrepancy tests. However the EST can be applied to sequences from any random number source, can be adapted to specific user requirements and has the added advantage that its computational complexity is relatively independent of the number of data points being tested. This later makes it particularly interesting as a test for extremely long period generators used for parallel computation work.

Using Geometrical Tools for Dimension Reduction of Images

❖Evgenia Rubinshtein, University of Central Arkansas, 2840 Dave Ward Dr, Apt J8, Conway, AR 72034, *evgeniar@uca.edu*; Anuj Srivastata, Florida State University

Key Words: Dimension reduction, image analysis, stochastic optimization, kurtosis, variance

We are interested in low-dimension linear representations of images, motivated the fact that image analysis often requires dimension reduction before statistical analysis, in order to apply sophisticated procedures. We present geometric tools for finding linear projections that optimize a given criterion for a given data set. We formulate this problem as multidimensional optimization on Stiefel manifold and use different criteria. The gradient vector is represented as the orthogonal projector onto the tangent space of the manifold. We use stochastic gradient methods to solve this problem in order to search for the global maximum. We demonstrate these results using several image datasets, including natural images and facial images.

151 Reducing Bias in Epidemiological Studies: Two-Phase Sampling and Missing Data Methods ●

Section on Statistics in Epidemiology, Section on Health Policy Statistics

Monday, July 30, 10:30 am–12:20 pm

A Method To Assess Bias Reduction in Observational Studies of Influenza Vaccine Effectiveness

❖Jennifer Nelson, Group Health Center for Health Studies, University of Washington, Department of Biostatistics, Seattle, WA 98101, *nelson.jl@ghc.org*; Michael Jackson, Group Health Center for Health Studies; Noel Weiss, University of Washington; Lisa Jackson, Group Health Center for Health Studies

Key Words: bias reduction, observational study, influenza, vaccine effectiveness

Accurately estimating the public health benefit of influenza vaccination for seniors is important for planning during pandemic and nonpandemic periods. Existing estimates of vaccine benefit in the elderly are primarily derived from observational studies. These studies typically use data from research or health care utilization databases and consistently report substantial reductions in all-cause mortality and respiratory disease hospitalization risk during influenza season in vaccinated seniors. However, there are reasons to believe that bias due to differences between vaccinated and unvaccinated seniors may account for some or all of these differences. We summarize the evidence for bias in published studies, propose design and analysis strategies that may yield more accurate estimation, and present a novel method to assess the ability of such strategies to reduce bias in future studies.

Multiple Imputation To Control for Unmeasured Confounding Using an Internal Validation Study

❖Robert Glynn, Brigham and Women's Hospital, 900 Commonwealth Avenue, Boston, MA 02215, *rglynn@rics.bwh.harvard.edu*; Til Stürmer, Brigham and Women's Hospital

Key Words: confounding, bias, regression models, multiple imputation, missing data, propensity scores

Multiple imputation (MI) can be used to adjust main study estimates for unmeasured confounding if additional information on confounders is available from a validation study. We compared MI of a missing covariate and of the gold-standard propensity score (PSG). We simulated cohort studies with a binary exposure A, a Poisson outcome Y, and 3 independent, normally distributed confounders, of which 1 (C) was only observed in a validation sample. MI of C used linear regression of C on A, Y, and measured covariates; whereas MI of PSG used linear regression of PSG on A, Y and PSE (the error-prone propensity score without C). Both approaches performed well to control for unmeasured confounding and led to substantial bias reductions which were slightly less pronounced with MI of PSG. However, MI of PSG is easily implemented with multiple missing confounders and might be useful in that setting.

Estimation of Causal Effects in Studies with Outcome-Dependent, Two-Phase Sampling

❖Weiwei Wang, Johns Hopkins Bloomberg School of Public Health, 615 N Wolfe St, E3040, Baltimore, MD 21205, *wewang@jhsph.edu*; Daniel Scharfstein, Johns Hopkins Bloomberg School of Public Health; Zhiqiang Tan, Johns Hopkins Bloomberg School of Public Health; Ellen MacKenzie, Johns Hopkins Bloomberg School of Public Health

Key Words: outcome dependent, two-phase, causal inference, biased sampling, semiparametric, doubly robust

We consider studies in which it is inexpensive to measure an outcome Y, a nonrandomized binary treatment T and a subset of confounding factors V, but expensive to measure additional confounding factors W. In such studies, outcome-dependent two-phase sampling can significantly reduce the cost. We propose two estimators of the causal effect of treatment: doubly robust and locally efficient, and compare them to the simple inverse weighted estimator. We illustrate our methods with data from the National Study on the Costs and Outcomes of Trauma and demonstrate finite sample performance of the estimators in a simulation study. We argue that the doubly robust estimator provides the best tradeoff in terms of robustness, efficiency and ease of implementation.

Analysis of Epidemiological Studies with an Outcome-Dependent Sampling Design

❖Haibo Zhou, The University of North Carolina at Chapel Hill, Department of Biostatistics, CB7420, Chapel Hill, NC 27599, *zhou@bios.unc.edu*

Key Words: outcome dependent sampling, continuous outcome, empirical likelihood, weighted analysis, epidemiologic methods, semiparametric

To characterize the relation between an exposure and a continuous outcome, the sampling of subjects can be done much as it is in a case-control study, such that the sample is enriched with subjects who are especially informative. In an outcome dependent sampling (ODS) design, observations made on a judiciously chosen subset of the base population can provide nearly the same statistical efficiency as observing the entire base population. Reaping the benefits of such sampling, however, requires use of an analysis that accounts for the ODS design. In this report, the authors examined the statistical efficiency of a plain random sample analyzed with standard methods, compared with that of data collected with an ODS design and analyzed by either of two appropriate methods.

Estimates of Causal Effects of a Continuous Exposure Using Multiple Imputation

❖Sundar Natarajan, Veterans Affairs New York Harbor Healthcare System/New York University, 423 E 23rd Street Room 11101S, New York, NY 10010, *natars01@med.nyu.edu*; Stuart R. Lipsitz, Brigham and Women's Hospital; Garrett Fitzmaurice, Harvard University; Scott Regenbogen, Harvard School of Public Health

Inferring causation is difficult in observational studies. Confounders may not be balanced across groups, and conclusions about 'exposure' effects are vulnerable to selection bias. For a categorical exposure, propensity score adjustments are popular. For a continuous exposure, these adjustments are difficult. Instead, we will use Rubin's multiple imputation approach to estimate the causal effects of a continuous exposure. Specifically, if there are G distinct exposure levels, for the i^th person, we create G-1 new observations with all possible confounders the same as the i^th person, but with the exposure replaced with the G-1 exposures that the subject didn't get; further, the outcome is set at missing. Multiple imputation is used to 'impute' missing outcomes. This approach is also used to estimate the mean outcome at 'doses' of the exposure of interest.

Nonparametric Regression with Missing Data Using Kernel-Estimating Equations

❖Lu Wang, Harvard University, 655 Huntington Ave, SPH2, 4th Floor, School of Public Health, Boston, MA 02115, *luwang@hsph.harvard.edu*; Xihong Lin, Harvard School of Public Health; Andrea Rotnitzky, Universidad Torcuato Di Tella

Key Words: nonparametric mean model, two-stage study, inverse probability weighted kernel GEE, augmented estimating equations, double robustness, asymptotic efficiency

In epidemiology studies when we do not know the functional pattern of how the mean outcome depends on a covariate, nonparametric regression becomes attractive. However, if nonresponse occurs nonrandomly (e.g., in stratified two-stage designs), naive approach using complete cases is usually biased. We propose inverse probability weighted kernel generalized estimating equations (GEE) and a class of augmented inverse probability weighted (AIPW) kernel GEE to correct for dependent censoring and nonrandom missing. Both approaches are asymptotically unbiased and do not require full likelihood specification. Further, the double robustness of AIPW kernel estimator provides us two chances to achieve valid inferences and it has potentials to gain efficiency. Simulations are done to evaluate practical performances, followed by an application on the AIDS Costs and Services Utilization Survey data.

Semiparametric Efficient Causal Inference with Missing Data

❖Yue Shentu, Merck & Co., Inc., 10 Landing Lane, Apt 6R, New Brunswick, NJ 08901, *yue_shentu@merck.com*; Minge Xie, Rutgers University

Key Words: Propensity Score, Observational Study, Missing at Random, Semiparametric efficiency bound

A missing data problem is investigated in the context of causal inference. In real-life observational studies, the response of interest may often be missing for a subgroup of subjects, and the missing mechanism may depend on post-baseline outcomes as well as baseline information. Motivated by a semiparametric efficiency bound of consistent treatment effect estimation, we proposed an estimator that incorporates the propensity-weighting and the regression imputations. We show that under mild assumptions, the proposed estimator is n-1/2-consistent and semiparametrically efficient. In addition, we show that the proposed estimator is robust against some of the model mis-specifications. Simulation studies were carried out to compare the proposed estimator with other existing estimators and to demonstrate its desirable properties.

152 Survey Measurement and Data Quality ●

Section on Survey Research Methods, Social Statistics Section

Monday, July 30, 10:30 am–12:20 pm

Can Study Design Influence Substantive Survey Results?

❖Dhuly Chowdhury, RTI International, 6110 Executive Blvd, Suite 902, Rockville, MD 20852, *dchowdhury@rti.org*; Karol Krotki, RTI International; Lily Trofimovich, RTI International

Key Words: measurement bias, field procedures

The objective of the Synar survey is to determine the buy rate for each state based on random, unannounced inspections of a probability-based sample of tobacco outlets accessible to youth. The buy rate is defined as the proportion of tobacco outlets that sell cigarettes to minors. Some researchers claim that states can change the buy rate by manipulating the age and gender distribution of the inspectors. We will explore to what extent inspectors' age and gender impacts the buy rate. Our primary hypotheses are 1) female inspectors are more likely to be able to buy cigarettes than male inspectors, and 2) older inspectors are more likely to be able to buy cigarettes than younger inspectors. We will also study interactions and the influence of covariates and will include procedures for the control of bias introduced by the study design.

SOI Develops Better Survey Questions Through Pretesting

❖Tara Wells, Internal Revenue Service, Statistics of Income, 500 North Capitol Street NW, Washington, DC 20001, *tara.r.wells@irs.gov*; Diane M. Milleville, Internal Revenue Service

Key Words: pre-testing, survey question development

Recently, the Statistics of Income (SOI) Division of the Internal Revenue Service helped a customer develop a survey. Without any prior knowledge of the survey's topic, SOI found it difficult to write "good" survey questions. Through the use of pretesting, SOI gradually became more familiar with the topic, determined how to phrase the questions, and understood which questions to include in the survey. Using cognitive interviewing, along with an Intranet application, SOI was able to obtain feedback from a small subset of the survey population. The survey content evolved into a set of well-developed questions that were easily understood by the participants. Through continued research of survey question development, SOI will benefit in future survey projects through the use of pretesting.

A New Approach to Measuring Residence Status

❖Jennifer Childs, U.S. Census Bureau, Statistical Research Division, 4700 Silver Hill Road, Washington, DC 20233, *jennifer.hunter.childs@census.gov*; Elizabeth M. Nichols, U.S. Census Bureau

Key Words: coverage, residency, split-panel field test

To determine where to count people who have multiple places to stay (e.g., second home, relatives' homes) on Census Day, Census coverage operations use a set of "cycle" questions that ask how often a person goes back and forth between places. These questions offer predefined patterns (e.g., going between places every week) and ask where the person spent most of the time during a specified time period (e.g., March and April). Because the cycle questions presume set patterns of living situations that may not reflect reality, we investigate an alternative method of assigning residency. The new "dates" method involves collecting dates of stays for each address and calculates where a person should be counted on census day given the dates. This paper presents findings from an experiment that evaluated the effectiveness of this alternative method of assigning residency.

Evaluating the Computer Audio-Recorded Interviewing (CARI) Household Wellness Study (HWS) Field Test

❖ Taniecea Arceneaux, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *taniecea.a.arceneaux@census.gov*

Key Words: Computer Audio-Recorded Interviewing, Quality Assurance, Computer Assisted Personal Interviewing

This report provides the results of research, testing, and analysis conducted by the U.S. Census Bureau. CARI will allow more efficient sampling methods and a more focused reinterview program while maintaining the Census Bureau's standard of data quality. With the respondent's consent, CARI unobtrusively recorded the verbal exchange between the interviewer and the respondent without disrupting the normal interview process. While the broader goal of this research is to implement CARI into all of the Census Bureau's computer-assisted personal interview (CAPI) surveys, the present study accomplishes the following tasks: evaluate the impact of CARI on data quality, measure the audio quality of recordings collected during personal visit interviews, determine the impact of CARI on system performance, and obtain the reactions of respondents Field Representatives (FRs) to the use of CARI.

Patterns of Nonresponse for Key Questions in NSDUH and Implications for Imputation

❖ Peter Frechtel, RTI International, One Metro Center, 701 13th Street NW Suite 750, Washington, DC 20005, *frechtel@rti.org*; Elizabeth Copello, RTI International

Key Words: imputation, nonresponse, nonresponse bias, probes, item nonresponse

The National Survey on Drug Use and Health (NSDUH), an annual nationwide survey involving approximately 67,500 subjects per year, does not make use of callbacks. However, for several key questions on drug use in the NSDUH, follow-up questions, or "probes," are presented to subjects who entered a response of "don't know" or "refused" to the original questions. The probes are a low-cost method of increasing item response rates, and of reducing nonresponse bias. An earlier study examined the success of the probes in increasing item response rates. This paper will mainly focus on nonresponse bias. There is strong evidence that respondents who answered the probes differ from respondents who answered the original questions, especially those who refused to answer the original questions. Some improvements to the current imputation method are proposed and evaluated.

The Impact on Data Quality of the Transition to Clean-Burning, On-Highway Diesel

❖ Paula Weir, Energy Information Administration, 19343 Mill Dam Place, Lansdowne, VA 20176, *paula.weir@eia.doe.gov*; Pedro J. Saavedra, Macro International; Benita O'Colmain, Macro International

Key Words: variance, bootstrap, on-highway, ULSD, price

The Environmental Protection Agency required that by June 2006 refiners and imports of petroleum must ensure that at least 80 percent of the volume of highway diesel fuel they supply be Ultra-Low Sulfur Diesel (ULSD). By December 2010, all on-highway diesel must be ULSD. Between 2006 and 2010, both ULSD and low sulfur diesel (LSD) may be sold at retail locations outside of California, with some outlets carrying both fuels and others carrying only one. Until January 2007, EIA collected the price of on-highway diesel without sulfur distinction. This paper describes how the weekly diesel price survey was modified to account for the transition to ULSD. One aspect central to the design was the estimation of the proportion of each kind of diesel sold in each sampling cell. An evaluation of the variance using a bootstrap method and sensitivity analysis of alternate assumptions are presented.

Interviewer Variability and Survey Estimates: Quantifying Data Quality

❖ Novie Younger, The University of the West Indies at Mona, Jamaica, P O Box 25580, KIN 3960, Miami, FL 33102-5580, *novie.younger@uwimona.edu.jm*; Damion Francis, University of the West Indies; Shelly McFarlane, University of the West Indies; Jan van den Broeck, University of the West Indies; Georgiana Gordon-Strachan, University of the West Indies; Andriene Grant, Ministry of Health, Jamaica; Deanna Ashley, Ministry of Health, Jamaica; Rainford Wilks, University of the West Indies

Key Words: Intra-observer correlation coefficient, observer design effects, observers, random effects models, precision, means

Ensuring data quality in national health surveys using large numbers of observers is important. Assessment of data quality and validity of survey estimates is desirable. Training and certification contribute to improved data quality but maintenance of quality in the field is often not assessed. Health survey data can quantify observers' influence on precision of sample means. Random effects models estimate association between observers' and supervisors' ("gold standard") independent replicate measurements adjusted for time lag. We estimate the intra-observer correlation coefficients (IOCC), observer design effects and the impact of observer variability on the variance of means. These results quantify survey data quality and suggest optimal intervals between interviewers' measurements and their "gold standard" for the purpose of quality control.

153 Survey Weighting and Calibration Methods ●

Section on Survey Research Methods

Monday, July 30, 10:30 am–12:20 pm

Comparing Estimate Bias and Variances Using Calibrated Weights with a Different Number of Constraints in a Sample with Oversamples in Specific Groups

❖ Nuria Diaz-Tena, TNS-Global, 22 Clearview Ave, Princeton, NJ 08540, *nuria.diaz-tena@tns-global.com*; Rob Cole, TNS-Global; Partha Dass, TNS-Global

Key Words: weight adjustment, weight calibration, marginal targets, web surveys, noncoverage, oversample

TNS conducts a web-based customer satisfaction for each of the major wireless service providers in 20 Designated Marketing Areas (DMAs). The survey measures both overall satisfaction as well as performance on a large number of attributes. The client needs to be able to look at the data in a variety of ways: the entire wireless market; by specific wireless provider; and particularly by specific wireless provider within individual DMAs. TNS computes weights to compensate for noncoverage, nonresponse and oversamples. To compensate for response bias, we decided to calibrate age, gender and income within DMA and provider. Previously, the calibration was done only by age, gender, income and provider within a DMA. Increasing the number of constraints to calibrate within area and provider improved the effective sample sizes of the DMAs by providers.

A New Algorithm for High-Dimensional Calibration in Observational Studies

❖ Michael E. Jones, Westat, 1650 Research Blvd, Rockville, MD 20850, *MikeJones@westat.com*; Ismael Flores Cervantes, Westat; David R. Judkins, Westat

Key Words: poststratification, raking, weighting

It has been shown recently that raking can be used as a balancing technique when applying propensity scoring to reduce bias in quasi experiments. When raking to balance, the potential number of control variables may increase to a level where time to convergence is impractical, even with high powered computers. As an alternative to the classic raking procedure, we describe a new method designed specifically for high-dimensional calibration. We evaluate the new method primarily in terms of convergence speed. However, we are mindful that this tool could facilitate questionable estimation strategies wherein people calibrate on many extraneous dimensions. To assess the consequences of such a strategy, we evaluate the performance of an estimator of a survey mean based on weights that were calibrated on many dimensions that are independent of the variable of interest.

Improving the Unit Nonresponse Adjustment in the NLSCY Using Logistic Regression Modeling and Calibration

❖ Mike Tam, Statistics Canada, Main Building 2500 K, 150 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *mike.tam@statcan.ca*; Marcelle Tremblay, Statistics Canada; Sarah Franklin, Statistics Canada; Claude Girard, Statistics Canada

Key Words: Nonresponse, Logistic Regression, Calibration

The National Longitudinal Survey of Children and Youth (NLSCY) is a longitudinal survey conducted by Statistics Canada. For the first five cycles, unit nonresponse for the original cohort of 0–11 year-olds was treated by reweighting within RHGs formed via segmentation modeling. In an effort to improve the longitudinal consistency of the estimates, the use of logistic regression to form the RHGs was investigated and implemented at cycle 6 of the NLSCY. Since the model that was used in production included only frame and cooperation variables, further development of the logistic regression model for cycle 7 of the NLSCY is planned. The potential improvements, including incorporating more questionnaire variables from all cycles and the use of paradata, are discussed. We also present the results of using calibration as another way to achieve longitudinal consistency.

Weight-Trimming in the National Immunization Survey (NIS)

❖ Sadeq Chowdhury, NORC at the University of Chicago, 1350 Connecticut Ave, NW, Suite 500, Washington, DC 20036, *sadeqc@yahoo.com*; Meena Khare, National Center for Health Statistics; Kirk Wolter, NORC at the University of Chicago

Key Words: Extreme weights, outliers, trimming, truncation

Excessively large sampling weights can unduly inflate variances of survey estimates. The NIS weighting procedure involves a series of nonresponse and undercoverage adjustments that introduce a considerable variation in sampling weights, even though the sample is selected with equal probability within each of 78 immunization action plan (IAP) areas. To control such undue variance in estimates, extreme weights are truncated and the remaining weights are adjusted to compensate for the truncation. We plan to present an analytical approach to assess the effects of weight trimming on the NIS estimates. The results will provide an indication whether there is scope for further optimization of the trimming procedure. Also, an alternative approach for detecting extreme weights and setting trimming cutoff values under the assumption of an exponential distribution of the weights will be discussed.

Testing for Informative Weights and Weights Truncation in Multivariate and Multilevel Modeling with Survey Data

❖ Tihomir Asparouhov, Muthen & Muthen, 2723 Sherwood Dr, Salt Lake City, UT 84108, *asparouhov@hotmail.com*; Bengt Muthen, University of California, Los Angeles

Key Words: Weights Ignorability, Weights Truncation, Survey data modeling, Multivariate Models, Multilevel Models

Analyzing the informativeness of the sampling weights can lead to significant improvement in the precision of model estimation with survey data. Testing for weights ignorability and weights truncation are two techniques that can be utilized in the sampling weights analysis. We evaluate the performance of these techniques in simulation studies based on multivariate factor analysis model as well as multilevel linear and logistic regression models. To test the weights ignorability we use Pfeffermann's (1993) test. We also compare different weights truncation techniques and discuss strategies for determining the optimal level of truncation. We describe the implementation of these techniques in the software package Mplus.

An Application of Alternative Weighting Matrix Collapsing Approaches for Improving Sample Estimates

❖ Linda Tompkins, Centers for Disease Control and Prevention, National Center for Health Statistics, 1810 Bunker Hill Road NE, Washington, DC 20018, *lit3@cdc.gov*; Jay J. Kim, National Center for Health Statistics

Key Words: Coverage ratio, Sample weighting, Cell collapsing

In sample weighting, most surveys combine small race groups such as American Indians, Asians, and other races with Whites. One problem with this approach is that Whites have a much better coverage ratio than the smaller race groups. This results in an overestimation of Whites and an underestimation of smaller race groups. For example, in weighting the 2003 National Health Interview Survey (NHIS) sample, the number of American Indians was underestimated by 29.70 percent and the number of Whites children under age year one were overestimated by 7 percent. The overall coverage ratio for Whites was 84.8 percent, while that for American Indians was only 50 percent. In this paper, alternative approaches for cell collapsing applied to the 2003 NHIS and the resulting estimates will be compared.

Bayesian Weight-Smoothing Models in Clustered or Cross-Classed Sample Designs

❖ Xiaobi Huang, University of Michigan, 2013 Medford Road, Apt H162, Ann Arbor, MI 48104, *xiaobih@umich.edu*; Michael R. Elliott, University of Michigan

Key Words: Sampling weights, clustered sample design, cross-classed sample designs, weight trimming, weight smoothing

Highly disproportional sample designs have large weights, which will introduce undesirable variability in statistical estimates. Weight trimming fixes a cutpoint weight and sets larger weights to this cutpoint value while adjusts weights below this value to maintain the untrimmed weight sum, reducing variability at the cost of introducing some bias. Previous work developed Bayesian "weight smoothing" models to produce general model-based weight trimming estimators of population statistics, but has been limited to the context of stratified and post-stratified sample designs. This presentation extends the Bayesian "weight smoothing" methodology to a more general class of complex sample design that include single or multi-stage cluster samples and/or strata that "cross" the weight strata. The methods are applied to linear and generalized linear regression models.

154 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, ENAR, WNAR, Section on Statistical Consulting, Section on Statistical Computing

Monday, July 30, 10:30 am–12:30 pm

A Quantitative Assessment of Diagnostic Cutoff Point Selection Methods

❖Michael Andrew, Centers for Disease Control and Prevention, 1095 Willowdale/NIOSH, Morgantown, WV 26505, *mandrew@cdc.gov*; Shengqiao Li, Centers for Disease Control and Prevention; Harshinder Singh, West Virginia University

Key Words: Diagnostic cutoff points, Specificity, Sensitivity, Predictive power positive, Predictive power negative, Prevalence

In cases where a decision about binary categorization of disease status is made on the basis of measurement on a trait, the determination of the cutoff point for the trait is an important problem. There are several criterion considered in literature for selecting the diagnostic cutoff point for classifying a subject in to disease or disease-free group. Notable among these are the probability of misclassification criterion, sum of sensitivity and specificity maximization criterion, maximal chi-squared criterion and a distance minimization criterion. We provide a quantitative assessment of different methods of obtaining diagnostic cutoff points in terms of performance of the resulting diagnostic quality measures such as sensitivity, specificity, predictive value positive and predictive value negative assuming models for the distribution of the underlying trait and the measurement error.

Application of Principal Component Analysis in the Reduction of a 19-Item Clinical Scale

❖Hong Liu-Seifert, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *liu-seifert_hong@lilly.com*

Key Words: Principal Component Analysis, factors, data reduction

Rating of Medication Influence (ROMI) is a 19-item clinical scale. Its use has been limited due to the large number of items and lack of established structure. The objective of this research was to apply Principal Component Analysis (PCA) to explore the underlying dimensions of ROMI. PCA was applied to ROMI from 2 studies with different designs and patient populations. Further, factors were derived at multiple time points and for each treatment arm for sensitivity. PCA yielded largely identical factors for ROMI for the 2 studies. These factors were consistent with the commonly accepted Health Belief Model. Results were also consistent at different time points and for each treatment group in each study. The derived factor structure for ROMI based on PCA is robust and may facilitate the use of ROMI as an easy and more applicable rating scale with reduced dimensions.

Statistical Applications in High-Throughput siRNA Synthetic Lethality Screens

❖Yue Li, Lilly Singapore Center for Drug Discovery, 1 Science Park Road 0401, The Capricorn, Singapore, 117528 Singapore, *liyue828@gmail.com*; Gopinath Ganji, Lilly Singapore Center for Drug Discovery; Holly Yin, Translational Genomics Research Institute; Quick Que, Translational Genomics Research Institute; Kok Long Ang, Lilly Singapore Center for Drug Discovery; Libin Ma, Lilly Singapore Center for Drug Discovery; Song Wu, Lilly Singapore Center for Drug Discovery; Greg Tucker-Kellogg, Lilly Singapore Center for Drug Discovery

Key Words: RNAi, HTS, Z prime factor, sigmoidal dose response model, IC50

RNA interference (RNAi) by small interfering RNA (siRNA) and short hairpin RNA (shRNA) reagents has emerged as a powerful technique to knock down the expression of specific genes. This has accelerated target validation efforts in drug discovery and development, particularly through cell-based phenotypic screening assays. An example of such an application involves "synthetic lethality" experiments in combination with small molecules for target or biomarker discovery. We have used this approach in a genome-wide high throughput screen (HTS) and examples of various statistical considerations and applications throughout the HTS, including experimental design, quality control and hit selection, will be presented and discussed.

Outlier Detection for a High-Throughput Bioassay

❖Shea Watrin, Amgen Inc., 4000 Nelson Rd, Longmont, CO 80503, *swatrin@amgen.com*; Ye Frank, Amgen Inc.

Key Words: Bioassay, Outliers

Bioassays are used to evaluate the potency of pharmaceuticals. These cell based assays can produce extreme results for which an assignable cause is not apparent. In a high throughput lab environment it is desirable to quickly and efficiently discern outliers from good data points. Investigating individual outliers is time consuming and can lead to inconsistent rejection of data. It is therefore desirable to set rules that will allow rejection of outliers automatically. These rules should protect against outliers but allow for variability that is common in the assay. We establish guidelines that allow the data to be evaluated for outlying data points before assumptions of the assay are verified and calculation of potency is carried out. These guidelines protect against subjective decisions about outliers and allow a high-throughput lab to process data efficiently.

The Effects of Sampling and Sample Preparation on Analytical Sensitivity

❖Thea Eudey, California State University, East Bay, Statistics Department, Hayward, CA 94542, *lynn.eudey@csueastbay.edu*; Richard Lewis, GlaxoSmithKline

Key Words: Analyte, Poisson distribution, Sensitivity

We investigate the probability of obtaining a positive assay result when the concentration of the analyte is low. In this situation, sampling and/or sample preparation can have significant effects on assay sensitivity, especially for assays that use small sample volumes. For example, due simply to chance, the analyte may not be present in a small sample, or, even if present in the sample, may not make it all the way through the sample preparation steps. In order to explore the potential impact of these effects, we use a relatively simple two-stage probability model. The first stage (sampling) employs a Poisson distribution and the second stage (sample preparation) employs a binomial distribution. We illustrate our findings with several examples.

Binary Logistic Regression Modeling of Precursor Sequence Cleavage

❖Allison Tegge, University of Illinois at Urbana-Champaign, 1207 Gregory Dr, Urbana, IL 61801, *ategge@uiuc.edu*; Sandra Rodriguez-Zas, University of Illinois at Urbana-Champaign; Bruce Southey, University of Illinois at Urbana-Champaign

Key Words: binary logistic model, bioinformatics, sequence analysis

Neuropeptides are signaling molecules critical in neural communication. Approaches to predict neuropeptides resulting from complex posttranslational enzymatic processing of precursors can support efficient experimental confirmation. We used a binary logistic model to predict precursor cleavage using amino acid location information. Logistic regression models were trained and cross-validated using precursor sequences and cleavage information from multiple mammalian species. Complementary variable selection methods were used to identify parsimonious models. The sensitivity of the models to correctly predict cleavages sites ranged from 68 to 75%. The species-specific logistic cleavage predictive equations differed in the explanatory variables and regression coefficients. Logistic modeling helped uncover species-specific sequence features influencing precursor sequence cleavage.

Frequentist and Bayesian Complement Each Other

❖Tie-Hua Ng, Food and Drug Administration, 1401 Rockville Pike, CBER, HFM219, Rockville, MD 20852, *tiehua.ng@fda.hhs.gov*

Key Words: Frequentist, Bayesian, sensitivity, positive predictive value, type I error rate, false discovery rate

A frequentist describes the data (or outcomes) conditioned on the unknown parameters while a Bayesian describes the parameters conditioned on the observed data. The two approaches are shown to complement each other in two applications, namely, (i) diagnostic test kit, and (ii) hypothesis testing. In the first application, the frequentist evaluates the sensitivity and specificity, while the Bayesian considers the positive and negative predictive values. In the second application, the frequentist controls the type I error rate while the Bayesian evaluates the false discovery rate. Recently, Ng (2007) argues that simultaneous testing for noninferiority and superiority increases the false discovery rate for superiority, although the type I error rate is controlled.

Sample-Size Determination for a Flow Cytometry Assay

❖Elaine A. McVey, BD Technologies, 21 Davis Drive, Research Triangle Park, NC , *Elaine_McVey@bd.com*; Perry Haaland, BD Technologies; Friedrich Hahn, BD Technologies

Key Words: flow cytometry, resampling, statistical computing, bioinformatics, sample size, detection limit

Flow cytometry is an important assay technology that has been limited by a lack of computational tools appropriate for use by statisticians and bioinformaticians. In this application, we demonstrate the use of a new open source software package to perform standard statistical analyses on flow cytometry data. In this case, cell numbers available for assay were limited, and sample size planning was required. Data was available from a calibration experiment which mimicked the expected experimental conditions and provided large numbers of observations. Using the flowCore package in R, these flow cytometry data were resampled to obtain proper statistical estimates of assay sensitivity parameters for small sample sizes.

Testing for Ordered Alternatives in the Analysis of Genetic Association Studies

❖Julia Kozlitina, Southern Methodist University, P.O. Box 750332, 3225 Daniels Ave, Dallas, TX 75275, *jkozliti@smu.edu*

Key Words: genetic, association, robust, order-restricted alternatives

Genome-wide association studies hold great potential for helping scientists understand the genetic component of complex human disease. However, the identification of individual variants underlying disease phenotypes remains a challenge for several reasons. Most common variants are likely to have only a minor effect on phenotype, while those alleles that do effect phenotype directly are likely to be found in populations at very low frequencies, so statistical tests will often suffer from low power to detect existing associations. We propose a robust procedure with higher power to detect association, especially when the underlying disease model is unknown.

Array Analysis To Assess Effects of RNAi-Based Silencing in Relation to Flavonoid Transport

❖Robert Norton, California State University, East Bay, 28022 El Portal Drive, Hayward, CA 94542-2512, *rnorton@horizon.csueastbay.edu*; Mousumi Rath, California State University, East Bay; Claudia Uhde-Stone, California State University, East Bay

Key Words: gene expression, array analysis, RNA, lupin

In many soils, low availability of phosphorus (P) is a limiting factor in crop cultivation. White Lupin (Lupinus albus L.) is a legume that has unique adaptations for extracting P from soil when other plants cannot. In response to P deficiency, white lupin forms densely clustered roots, referred to as proteoid (or cluster) roots. A previous cDNA array approach identified genes up-regulated in proteoid roots under P deficiency, including LaMATE (Lupinus albus multi drug and toxin efflux). The encoded protein displays homology to the MATE protein family of transporters. LaMATE has been shown to be also induced under other nutrient stresses. Furthermore, RNAi-based gene silencing was shown to effectively reduce LaMATE expression in transformed white lupin roots. Here, we are using nylon filter arrays to analyze gene expression in transgenic lupin roots silenced for LaMATE expression.

A Semiparametric Mixture Model for Identifying Cis-Acting Regulatory Elements

❖Gregory Hather, University of California, Berkeley, 367 Evans Hall, Department of Statistics, Berkeley, CA 94720-3860, *ghather@berkeley.edu*; Terence Speed, University of California, Berkeley; Mary Wildermuth, University of California, Berkeley

Key Words: cis-acting regulatory elements, binding sites, transcriptional regulation, sequence analysis

Traditional computational methods for identifying cis-acting regulatory elements do not effectively integrate gene expression data into the analysis. We present a model based on the work of Wang et al. (2005) that relates these two types of data. One input to our model is a list of short DNA segments that fall within the promoter region of a gene and that match a user-specified core motif. The other input is the gene expression data. Each DNA segment is assumed to be either active or inactive in terms of its role in regulating gene expression. Any gene with at least one active segment nearby is assumed to be active. Activity is a binary hidden variable that depends upon the sequence data and that influences the expression data. The model can be used to estimate the probability that a given DNA segment is active under the conditions for which the gene expression was measured.

Reproducibility of Microarray Data: A Further Analysis of Microarray Quality Control Data

❖Chien-Ju Lin, National Center for Toxicological Research, FDA, 3900 NCTR RD, Jefferson, AR 72079, *chien-ju.lin@fda.hhs.gov*; James J. Chen, Food and Drug Administration; Huey-Miin Hsueh, National Chengchi University; Robert R. Delongchamp, National Center for Toxicological Research; Chen-An Tsai, Academia Sinica

Key Words: interplatform comparability, microarray platform reproducibility, titration mixture, gene selection

The MicroArray Quality Control (MAQC) data provide an opportunity to assess reproducibility of gene expression data across multiple sites and multiple platforms. We evaluate 12091 common genes for four titration samples generated from five microarray platforms tested in three sites with five technical replicates. A total of 293 arrays is analyzed in the intra- and inter-platform comparisons. Hierarchical cluster analysis and analysis of variance models are used to assess reproducibility and consistency of the measured intensities. There are differences in intensities measured among the five platforms. Within each platform, all five platforms can well distinguish the four samples, but there are strong site effects. The four titration samples follow the concentration response relationship.

An EM Algorithm for Identifying Genotypic Structure Using Genome-Wide Expression Data

❖Ellen Breazel, University of Georgia, 3 Pittler Drive, Greenville, SC 29607, *ellenh@uga.edu*; Paul Schliekelman, University of Georgia

Key Words: complex traits, EM algorithm, genotypic structure

Recent studies have found that genomewide expression data may be a useful tool in the difficult task of mapping complex traits. We have developed a method using expression level information to cluster individuals by their genotype on disease causative loci. Standard clustering methods are not well suited to identify genotypic structure because they tend to be overwhelmed by variation that is unrelated to disease genetic variation. We propose an EM algorithm-based method that targets the disease genetic variation and will thus identify disease genotypic variation via the correlation structure in differences in gene expression between disease affected and unaffected individuals. Identifying genotypic structure in a population will allow gene mapping studies to take into account heterogeneity in disease genotype and will improve mapping power.

A Novel Application of Quantile Regression for Identification of Cartilage Biomarkers in Equine Microarray Data

❖Liping Huang, University of Kentucky, Lexington, KY , *liping@ms.uky.edu*; Christopher Saunders, George Mason University; Wenying Zhu, University of Kentucky; Jamie MacLeod, University of Kentucky; Arnold Stromberg, University of Kentucky; Arne Bathke, University of Kentucky

Key Words: quantile regression, microarray data

Total RNA from equine articular cartilage and 10 other tissues was isolated. Two color array experiments were performed to compare the gene expression profile of cartilage versus each of the other 10 tissues. Cartilage-specificity of expression was determined using quantile regression analysis by estimating conditional quantiles of the log2 ratios given the observed average log2 intensities. Nine conditional quantiles were estimated. For each observed gene intensity in a given tissue comparison, a cartilage-specificity score was generated by substituting the Z-score for the corresponding estimated regression quantile. Thirty-eight probe sets were identified that exhibit expression above the 95th conditional quantile in all 10 of the cartilage/tissue comparisons analyzed. The list includes established biomarkers of articular cartilage.

A Novel Method of Identifying Differentially Expressed Genes Based on Probe-Level Data for GeneChip Arrays

❖Zhongxue Chen, Southern Methodist University, 3225 Daniels Ave Statistical Science Dep, POBox 750332 SMU, Dallas, TX 75275, *zhongxue@smu.edu*; Monnie McGee, Southern Methodist University; Qingzhong Liu, New Mexico Institute of Mining and Technology; Megan Kong, The University of Texas Southwestern Medical Center at Dallas; Richard Scheuermann, The University of Texas Southwestern Medical Center at Dallas

Key Words: Differentially Expressed Gene, Microarray, False Positive Rate, Receiver Operating Characteristic, spike-in, GeneChip

Current gene selection methods suffer from the -multiple comparisons problem due to the large number of significance tests that must be performed. Methods that control false discovery rate (FDR) are difficult to estimate the number of true negatives and are limited to the application for real data. We propose a novel method for identifying differentially expressed genes (DEGs), probe level identification of differentially expressed genes (PLIDEG). With the extra information provided by probe level data, PLIDEG can not only control type I error but also increase the power of detecting DEGs simultaneously. Therefore, PLIDEG can efficiently separate differentially expressed genes and non-DEGs without requiring estimation of the number of non-DEGs. Based on theoretical analysis and real microarray data, we show that PLIDEG has better performance than other methods.

A Network-Based Gene Screening Approach

❖Wei Zhao, University of California, Los Angeles, Department of Human Genetics, Los Angeles, CA 90095, *wzhao@mednet.ucla.edu*; Paul Mischel, University of California, Los Angeles; Steve Horvath, University of California, Los Angeles

Key Words: Network, Microarray, Gene Screening

Weighted gene co-expression network analysis (WGCNA) has been shown successful in selecting cancer relevant genes using microarray data. We systematically study the WGCNA based gene screening method and compared it with traditional gene screening methods. Using public available microarray data sets, we show that the genes with high connectivity in clinically relevant modules are strikingly and reproducibly associated with the clinical outcome. These results demonstrate that gene co-expression networks contain valuable information for identifying prognostically important individual genes. Our simulations and empirical studies show that network connectivity can be an important additional gene screening variable for selecting biologically and statistically significant genes.

Statistical Methods for Peptide Identification Using Tandem Mass Spectrometry Through Sequence Database Searching

❖Qunhua Li, University of Washington, 4747 30th Ave NE 155, Seattle, WA 98105, *qli@stat.washington.edu*; Matthew Stephens, The University of Chicago

Key Words: mixture models, EM algorithm, mass spectrometry, proteomics, peptide identification

Peptide identification using tandem mass spectra through database searching is a key component in proteomic research. In this study, we developed a model-based scoring algorithm for peptide identification using mixture models. It effectively takes account of the spectral noise and incorporates the spectral intensity. In addition, our approach also provides a natural confidence measure for assessing the uncertainty of the score. We will present our results on several public available benchmark datasets.

Development and Application of a Multivariate Mantel-Haenszel Mean Scores Test

❖Rebecca McNeil, Medical University of South Carolina, Department of Biometry, 135 Cannon St Suite 303, Charleston, SC 29425, *boehm@musc.edu*; Robert Woolson, Medical University of South Carolina; Carol Wagner, Medical University of South Carolina

Key Words: Mantel-Haenszel, multivariate, stratified, clustered, mean scores test, clinical trial

Several methods are appropriate for data from clinical trials with stratified randomization: these include linear models and the Mantel-Haenszel mean scores test. It has been demonstrated that the mean scores test has superior power and size to the linear model alternatives in univariate data. Thus, we provide the theoretical foundations for a convenient multivariate extension to this highly useful test. In a series of Monte Carlo simulations of the comparative behavior of the Mantel-Haenszel and two linear models approaches in multivariate data, we found that the Mantel-Haenszel extension performed as well or better than the linear models in a wide range of data structures. We expect that this randomization model approach to multivariate or clustered stratified data will be a useful addition to the statistician's toolbox.

Curtailed Two-Stage Design for Phase II Clinical Trial

Yunchan Chi, National Cheng Kung University; ❖ Chiamin Chen, National Cheng-Kung University, Department of Statistics, *cmchen@email.stat.ncku.edu.tw*

Key Words: inverse binomial sampling, curtailed two-stage design, expected sample size, Simon's design

In general, patients enter a Phase II clinical trial sequentially and hence their responses to the treatment are reported consecutively. A long observational period is required, when the accrual rate is low and the treatment period is long. Simon's two-stage designs (1989) have recently employed to avoid giving patient an ineffective drug. Thus if the new drug is ineffective then this design definitely would shorten the drug development process. However, for a promising new drug this design may still require a long observation period. Therefore, this paper proposes an alternative curtailed two-stage design to shorten drug development process as soon as the treatment shows lack of efficacy or very effective. The savings in patient accrual compared to Simon's designs is about 5% and 40% for very low and high null response rates, respectively.

Statistical Evaluation of Bone Mineral Density as a Surrogate Endpoint for Fracture with or Without Baseline Consideration

❖ Peiqi Chen, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *chenpe@lilly.com*; Asad Rana, Eli Lilly and Company

Key Words: Bone mineral density, surrogate endpoint, fracture risk

The statistical methodologies to validate bone mineral density (BMD) change as a surrogate for fracture risk reduction are still in debate. The objective of this paper was to determine whether BMD change or the endpoint BMD value is the more appropriate covariate for fracture risk. Using data from a large osteoporosis trial, we examined (1) the relationship between BMD change and fracture risk; (2) the relationship of baseline BMD and BMD change with fracture risk; and (3) the relationship between the endpoint BMD value and fracture risk. The unintuitive relationship observed was that a high fracture risk was associated with a large increase in BMD when we ignored the baseline value. We further found that both the baseline BMD value and BMD change were important determinants of fracture risk. Therefore, it is more appropriate to use the endpoint BMD value in statistical modeling.

A Composite Endpoint for Vulvovaginal Symptoms in VVA Clinical Trials

❖ Moh-Jee Ng, Food and Drug Administration, 3823 Park Lake Dr, Rockville, MD 20853, *mohjee.ng@fda.hhs.gov*; Ling Chen, Food and Drug Administration

Key Words: menopausal women, Composite Endpoints, Vulvar and Vaginal Atrophy

Vulvar and vaginal atrophy (VVA) is an indication associated with menopausal women. Based on the draft Guidance for Industry, there are three co-primary endpoints for efficacy evaluation. One of them is mean change from baseline to week 12 in the moderate to severe symptom that has been identified by the patient as being the most bothersome to her. This endpoint is determined using five vulvovaginal symptoms which are included in the patient self-assessment questionnaires. It has been observed that some patients missed to identify this endpoint at the baseline, and some patients had more than one symptom mostly bothersome to them. In this research, we will propose a composite endpoint for evaluating vulvovaginal symptoms. Real data results will be presented in this poster.

Predicting Accrual in Clinical Trials: Bayesian Posterior Predictive Distribution

❖ Stephen D. Simon, Children's Mercy Hospital, 2401 Gillham Road, Kansas City, MO 64108, *ssimon@cmh.edu*; Byron Gajewski, The University of Kansas Medical Center

Key Words: prior elicitation, exponential, inverse gamma, Bayesian, sample size

Investigators need good statistical tools for the initial planning and the ongoing monitoring of clinical trials. In particular, they need to carefully consider the accrual rate-how rapidly patients are being recruited into the clinical trial. A slow accrual decreases the likelihood that the research will provide results at the end of the trial with sufficient precision to make meaningful scientific inferences. In this paper we present a method for predicting accrual across a fixed period of time. Using a Bayesian framework we combine prior information with the information known up to a monitoring point to obtain a prediction. We provide posterior predictive distributions of the accrual. The approach is attractive since it accounts for both epistemic and aleatory uncertainties. We illustrate the approach using actual accrual data.

Inference in Clinical Trial Using ROC Curve Regression

❖ Lin Zhang, Baylor University, Department of Statistical Science, One Bear Place 97140, Waco, TX 76798, *linzhg78@yahoo.com*; Jack Tubbs, Baylor University

Key Words: nonparametric, clinical trial, interaction effect, ROC curves regression, bootstrap

This paper considers a method proposed by Dodd and Pepe where the area under the ROC curve (AUC) provides a measure of the treatment effect that allows for the incorporation of strata variables, covariates, and their possible interaction in the usual logistic regression model. Dodd and Pepe proposed using bootstrapped standard errors as the independence assumption of the logistic model was violated when using a function of the usual Mann-Whitney statistic as the response variable. In this paper we consider the 'exact' estimates for the variance of the estimate of the AUC as a function of strata variables and discrete covariates. The confidence interval for the AUC based on this method is compared with the bootstrap procedure.

Semiparametric Smoothing for Scintigraphy Data from a Three-Period Crossover Design

❖ Jo Edmonds, Baylor University, 9821 Chapel Road, Apartment 1621, Woodway, TX 76712, *jo_edmonds@baylor.edu*; John W. Seaman, Jr., Baylor University

Key Words: semiparametric regression, penalized spline models, smoothing

We consider semiparametric regression as a smoothing technique in studying the relationship between gastric emptying and time. Our analysis utilizes scintigraphy data from a three-period crossover design, with subjects randomized to one of three sequences of dose administration. The relationship is modeled with a semiparametric smooth function, specifically a penalized spline model, implemented as a linear mixed model. Estimates of the smooth functions are obtained through the use of SAS PROC MIXED. Separate models are obtained for each of three dose levels for two types of meals.

A Performance Analysis of Distribution-Free Tests for Drug Interaction

❖ Yuhong Tang, Sam Houston State University, 913 University Ave, Apt 12, Huntsville, TX 77320, *yxt001@shsu.edu*; Melinda M. Holt, Sam Houston State University

Key Words: Bootstrap, Likelihood Ratio Test, Synergy, Power, Type I Error, Clinical Trials

One prevalent goal within clinical trials is to determine whether or not two drugs interact when applied in combination. Holt, Stamey, Seaman, and Young (2004) developed the likelihood ratio test and proposed a new bootstrap test for drug interaction, assuming that non-interaction is modeled by statistical independence. The methods discussed by Holt, Stamey and Young were specifically targeted to fixed-dose combination studies that involve binary response data. The authors provided a limited comparison of the Type I error and power of the two procedures using Monte Carlo simulation. We will present an expansion of that comparison and provide advice on choosing between the likelihood ratio and bootstrap tests in this context. We consider issues of small samples, unequal sample size, and bias.

Drug Combination Studies in Preclinical Pharmaceutical Research

❖Joseph Haas, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *haasjv@lilly.com*; Philip Iversen, Eli Lilly and Company

Key Words: drug combination, synergy, additivity, pharmaceutical research, pre-clinical

Drug combination studies are important in pre-clinical pharmaceutical research. If a combination is synergistic or additive, lower doses can be used to achieve the same or better efficacy with lower toxicity. We compare different definitions of additivity in the bio-medical literature, including Loewe additivity and Bliss independence. We also describe some experimental designs and recent dose-response analysis methods for combination studies. Real data examples will be shown from the fields of oncology and insomnia.

Evaluation of Soluble KIT as a Potential Surrogate Marker for Time-to-Tumor Progression in Patients with Advanced GIST by Several Recently Proposed Methods

❖Xin Huang, Pfizer Inc., 10555 Science Center Drive, San Diego, CA 92121, *xin.huang@pfizer.com*; Sam DePrimo, Pfizer Inc.; Jenny Zhang, Harvard University; Charles Harmon, Pfizer Inc.; John Smeraglia, Pfizer Inc.; Chuck Baum, Pfizer Inc.; Randy Allred, Pfizer Inc.

Key Words: biomarker, clinical trial, surrogate

Researchers are interested in exploring biomarkers that may act as surrogate endpoints for the primary clinical endpoints in clinical trials. A considerable number of statistical methods have been proposed for this topic recently. Sunitinib is a multitargeted tyrosine kinase inhibitor of KIT, PDGFRs, VEGFRs, RET and FLT3, approved for the treatment of imatinib-resistant/-intolerant gastrointestinal stromal tumor (GIST). Preliminary analysis in a sunitinib Phase I/II GIST study suggests that a decline in plasma soluble KIT (sKIT) levels correlates with clinical benefit. In this paper, we evaluated the potential of sKIT as a surrogate endpoint for time-to-tumor progression (TTP) using data obtained in a randomized, double-blind, placebo-controlled phase III study of sunitinib in patients with imatinib-resistant/-intolerant GIST by several recently proposed methods.

Exact Inference for Nominal and Rating Scale Data with Repeated Measurements in Clinical Trials with Parallel Group Design

❖James Lee, Daiichi Sankyo Pharma Development, 399 Thornall St., Edison, NJ 08837, *jlee@dsus.com*; John Dar Shong Hwang, B.R.S.I.

Key Words: Exact inference, nominal data, treatment*time interaction, treatment effect, conditional distribution, rating scale

For a clinical trial with two treatment arms having repeated measurements of nominal data as the response, this paper proposes a model-free exact inference on hypotheses commonly encountered in practice. The method is an extension of the binary data case previously reported. Rating scale data is considered as a transformation of the nominal data. Construction and justification of parameters which may be used to represent Treatment by Time Interaction, common Treatment Effect, common Time Effect for the data with 3 categories of response is developed. Existence and actual derivation of a conditional distribution which depends on desired parameters and is devoid of nuisance parameters is illustrated using two repeated time points. Generally the distributions may be simplified to become product of basic binomial or hypergeometric distributions.

Identification of Treatment Effects in the Presence of Deviation from Randomized Assignment: Application to a Surgical Setting

❖Corwin Zigler, University of California, Los Angeles, 11610 Iowa Ave, Apt 6, Los Angeles, CA 90025, *zigler@ucla.edu*

Key Words: instrumental variables, noncompliance, latent class model

Often in medical settings, adherence to randomized assignment of competing treatments proves unethical or simply unrealistic. Without advantages of randomization, standard approaches may not allow useful comparison of treatment effects. An instrumental variables framework has emerged as a useful tool for identifying treatment effects in settings where treatment received differs from treatment assigned (Angrist, Imbens, and Rubin 1996, JASA and Imbens and Rubin 1997, Annals of Statistics). We apply this framework in an example involving surgical versus non-surgical treatment of mandibular fractures in which clinicians were allowed to override random treatment assignment. We use Bayesian methods that utilize data characterizing patient injury to estimate the effect surgical treatment on length of hospital stay.

Semisupervised Technique and Its Learning Rate in Intelligent Systems

❖Morteza Marzjarani, Saginaw Valley State University, Computer Science Department, 7400 Bay Road, University Center, MI 48710, *marzjara@svsu.edu*

Key Words: Artificial Intelligence, Machine Learning

In the field of machine learning which is a branch of Artificial Intelligence (AI), semi-supervised learning (SSL) occupies the middle ground, between supervised learning (where all training examples are labeled) and unsupervised learning (where no label data are given). Interest in SSL has particularly increased in recent years due to the availability of data in variety of forms such as images, text, and bioinformatics. In this study, we will present SSL in depth, develop the algorithm/software, study the learning rate in such system, and than compare the efficiency of SSL to that of supervised, and unsupervised systems.

The Multinode Topological Overlap Measure for Gene Neighborhood Analysis

❖Ai Li, University of California, Los Angeles, 500 Landfair Ave, Los Angeles, CA 90024, *aili@ucla.edu*; Steve Horvath, University of California, Los Angeles

Key Words: gene co-expression network, neighborhood analysis, toplogiclal overlap measure

A pair of nodes in a network is said to have high topological overlap if they are both strongly connected to the same group of nodes. In gene networks, genes with high topological overlap have been found to have an increased chance of being part of the same biological pathway. The use of topological overlap serves as a filter to exclude spurious or isolated connections. A topological overlap measure can be used in network analysis for finding similar nodes or for defining a neighborhood of an initial set of nodes. The standard topological overlap measure allows one to measure pairwise topological overlap in an undirected network. Here we generalize this overlap measure to multiple nodes. The resulting multiTOM measure can be used to define a neighborhood of a initial set of nodes. We provide several applications that show the use of these neighborhoods.

Modeling Movement and Diving Behavior of Satellite-Tagged Sea Turtles

❖Eric Nordmoe, Kalamazoo College, 1200 Academy St, Kalamazoo, MI 49006, *enordmoe@kzoo.edu*; Cheong Hoong Diong, Nanyang Technological University

Key Words: modeling, dive patterns, longitudinal data, satellite-relayed data loggers, movement, sea turtles

Despite intense study, little is known about the movement and open sea diving patterns of sea turtles raised in captivity. Only recently, satellite-relayed tracking data has allowed biologists to track movements and diving behavior of turtles in the open sea. This poster presents the results of a study of twelve sea turtles fitted with satellite-tracking devices and released in August 2006 into international waters in the South China Sea. Our analysis uses statistical modeling techniques for longitudinal data to analyze and describe the movement and dive data collected by the satellite-relayed data loggers. We provide implications for ecological policy and direction for future research.

Repeated Classification as a Cost-Effective Sample Design To Test Association When There Are Random Misclassification Errors

❖Nathan Tintle, Hope College, 27 Graves Place, Hope College Math Dept, Holland, MI 49423, *tintle@hope.edu*

Key Words: misclassification error, genotyping error, replicate sample, repeat sample, experimental design

Misclassification errors are well-known to cause significant bias and loss of power in related hypothesis tests. While random misclassification errors have received a fair bit of attention in the literature, repeated classification to address these errors has not been as fully explored. Recently, a probabilistic model for repeated classification was proposed, along with a method for incorporating repeated classification data into a traditional chi-squared test of association between two categorical variables. We present a method for comparing a traditional sampling design (all units classified once) with repeated classification (some portion of the sample is classified multiple times) assuming a fixed budget for sampling and classification costs. We find that repeated classification can be cost-effective (i.e. have increased power) to test association in some situations.

Multivariate Exploration and Analysis of Crop Yield Components for Optimal Selection in Honeycomb Designs

❖Andy Mauromoustakos, University of Arkansas, 101 Agriculture Annex, AGRX 101, Fayetteville, AR 72701, *amauro@uark.edu*; Vasilia Fasoula, The University of Georgia; Kevin Thompson, University of Arkansas

Key Words: Honeycomb Designs, Parallel Plot, Spatial Designs, MANOVA, PCA, JMP

Honeycomb designs (HD) are a set of systematic designs capable of handling a large number of genetic entries and a large number of replications. The designs sample effectively for environmental diversity by means of large number of moving replicates. There are recently receiving attention in geostatistics literature as a preferred alternative to the popular rectangular grid designs by being more efficient especially in situations with the semivariogram model has a finite range. In the plant breeding literature these designs were developed to carry out efficient selection among genetic entries through the partition of crop yield into three genetic components and efficient selection of the best plants within the selected entries by means of the moving-ring single-plant selection. This paper will focus on the multivariate exploration and analysis of three popular yield components.

An Additive Model for a Heavily Right-Skewed Outcome

❖Amanda Allshouse, RTI Health Solutions, PO Box 12194, Research Triangle Park, NC 27709, *aallshouse@rti.org*; Jianmin Wang, RTI Health Solutions; Jasmine Mathias, RTI Health Solutions; William Irish, RTI Health Solutions

Key Words: Cost model, healthcare cost modeling, Skewed outcomes, health econometrics

When estimating and modeling healthcare costs, challenges arise rooted in the heavily right-skewed nature of the distribution of non-zero annual costs. No consensus currently exists regarding the most appropriate method for analyzing heavily right-skewed data bounded at zero. Traditional approaches to this problem include fitting the logarithmic-transformed data with an ordinary least squares regression, resulting in a multiplicative model when an additive interpretation is desired. An additive model to accommodate skewness and heterogeneity has been developed. Heavily right-skewed non-zero data were generated to emulate health-care costs, and the statistical properties of the additive model compared to the traditional approach were evaluated using simulation techniques.

A Study on Bayesian Signaling of a Prespecified Drug-Adverse Event Association in Post-Marketing Safety Studies

❖Yu-te Wu, Food and Drug Administration, 10903 New Hampshire Ave, Rm 5223, Silver Spring, MD 20993, *yu-te.wu@fda.hhs.gov*; George Rochester, Food and Drug Administration; Yi Tsong, Food and Drug Administration

Key Words: Post-marketing safety, Bayesian method, Rare adverse event, Epidemiology

Post-marketing safety studies are increasingly used to monitor the potential risk of specific drug-induced adverse events (AEs) and to resolve uncertainties not addressed in new drug applications. Currently popular study designs include active surveillance studies, epidemiologic studies and registries. These studies are primarily designed using the frequentist approach, which may require a prohibitively large sample size and does not allow for continuous monitoring. A Bayesian design may allow for earlier detection of the safety signal because of its flexibility in continuous monitoring. In this presentation, we will apply the Bayesian approach to the monitoring and detection of rare AEs in post-marketing epidemiologic safety studies to investigate the feasibility of the Bayesian design, and also evaluate how it performs compared to conventional approaches.

Comparing Exact and Asymptotic Confidence Intervals for a Function of Variance Components

❖Brent Burch, Northern Arizona University, PO Box 5717, Flagstaff, AZ 86011-5717, *brent.burch@nau.edu*

Key Words: Heritability, Linear mixed models, Pivotal quantity, REML estimation

Heritability, the extent to which a physical characteristic is passed from one generation to the next, is often equated to a function of variance components in mixed linear models having two variances. Using normality assumptions, one can compute exact confidence intervals for heritability.

These intervals differ from those based on the asymptotic distribution of the REML estimator. In this poster, exact and asymptotic intervals are constructed for models ranging from the one-way random effects model to animal breeding models having correlated observations.

Testing for Slope Equivalency of Two Orthogonal Analytical Methods

❖ Richard Burdick, Amgen Inc., 4000 Nelson Road, Mail Stop AC22A, Longmont, CO 80503, *rburdick@amgen.com*; Shea Watrin, Amgen Inc.; Michael Stoner, Amgen Inc.

Key Words: Analysis of covariance, Equivalence test, Generalized confidence interval

Statistical equivalence testing is used in the biopharmaceutical industry to support method transfer, method remediation, and process comparability. In this study, equivalence testing is used to demonstrate that two orthogonal methods, size exclusion (SE) chromatography and sedimentation velocity analytical ultracentrifugation (SV-AUC), are similarly capable of detecting/quantifying protein aggregates (i.e., that they respond equivalently to changes in the amount of analyte). The two techniques were run side-by-side on samples containing protein aggregates spiked in at six different levels. In this paper, we compare definitions of equivalence based on both the ratio and difference of slopes. Methods for determining definitions of equivalence are proposed and discussed. The ratio and difference methods are compared, and recommendations for selecting an appropriate method are provided.

Step-Down Pairwise Comparisons Based on the Studentized Range Distribution

❖ Scott Richter, The University of North Carolina at Greensboro, Mathematics and Statistics, Greensboro, NC 27402, *scott_richter@uncg.edu*; Melinda McCann, Oklahoma State University

Key Words: multiple comparisons, step-down, pairwise

A step-down testing procedure, based on Tukey's multiple pairwise testing procedure is proposed. The proposed procedure is more powerful than Tukey's procedure while still maintaining strong control of the familywise error rate. A simulation study shows that the new procedure is more powerful than a similar technique based on Bonferroni's adjustment, and compares favorably to the multiple range procedure of Ryan, Einot, Gabriel & Welch. A computer is recommended to implement the full step-down method. However, a simpler version using only two steps that is only slightly less powerful is presented. The two-step test can be carried out using an existing table of the studentized range statistic.

An Iterative Least Squares Process To Obtain Unbiased Variance Estimates at the Second Stage of a Two-Stage Model for Longitudinal Data

❖ Mary Bartholomew, Food and Drug Administration, 7500 Stanish Place, Center for Vet Medicine, Rockville, MD 20855, *mary.bartholomew@fda.hhs.gov*; Chris Gennings, Virginia Commonwealth University

Key Words: bias, Nelder-Mead algorithm, uncertainty, subject-specific, population-averaged

Two-stage models for longitudinal data are intuitively appealing because the modeler obtains a good picture of subject-specific models and the parameters of the population-averaged model. However, it was widely publicized that variance estimates for the population parameters based on variation among estimated subject-specific parameters were biased due to failure to account for uncertainty in the subject-specific estimates. Likelihood methods that correctly estimate the variance in the population model have been developed but the population-averaged parameters are obtained directly and often the subject-specific models are neglected when this approach is used. An iterative procedure using the Nelder-Mead algorithm to cycle between subject-specific estimates and population parameter estimates removes bias and was applied to well known orange tree growth data with good results.

Type I and Type II Error Properties for White's Robust Covariance Matrix Estimator in Longitudinal Designs with Several Repeats

❖ Keith Williams, University of Arkansas for Medical Sciences, 4301 W Markham Street, Slot 781, Little Rock, AR 72205, *williamsdavidk@uams.edu*; Deborah W. Biddle, University of Arkansas for Medical Sciences

Key Words: Longitudinal, Robust Covariance Estimation, Repeated Measures, White's Sandwich Estimator

SAS Proc Mixed and Genmod offer an empirical option to estimate the variance-covariance structure for repeated measure designs. The empirical option uses an estimation approach that is referred to as 'White's sandwich estimator,' and is regarded to possess robustness and statistical accuracy. Its use is generally reserved for settings in which the researcher does not have solid knowledge of a particular covariance structure that is appropriate for a data set at hand. This study investigates the Type I and II error performance of White's estimator when compared to competing approaches in various repeated measure design scenarios.

Effect of Ignoring Baseline on Transitions from Intact Cognition to Dementia

❖ Richard J. Kryscio, University of Kentucky, Lexington, KY 40505, *kryscio@email.uky.edu*; Lei Yu, University of Kentucky; David Snowdon, University of Kentucky; Suzanne Tyas, University of Waterloo

Key Words: dementia, mild cognitive impairment, Markov chain, baseline, random effect, polytomous logistic

We investigate bias introduced by ignoring baseline on transitions from intact cognition to dementia with mild cognitive impairment and global impairments as intervening states. Transitions among states are modeled by a Markov chain (MC) with transition probabilities depending on a single covariate. Two cases are examined by simulation (i) a true MC is left truncated and (ii) cognitive decline is assumed to be a linear model or a double change point model with random slopes. Results are illustrated with an application from the Nun Study, a longitudinal cohort of 678 participants 75+ years of age at baseline.

A Simulation Study Comparing Linear Mixed Models and Generalized Estimation Equations for Use in Longitudinal Biomarker Study for Exposure to Cigarette Smoke

❖ Qiwei Liang, Philip Morris USA, PO Box 26583, Richmond, VA 23261, *qiwei.liang@pmusa.com*; Huajiang Li, Quintiles Inc; Hans Roethig, Philip Morris USA

Key Words: Monte Carlo simulation, linear mixed models, generalized estimation equations, biomarkers, Type I error, cigarette smoke

The objective of this study is to use Monte Carlo simulation to compare linear mixed models and generalized estimation equations for application in clinical studies investigating human exposure to cigarette smoke. Unlike some other simulation studies for model comparisons, we specify the simulation study design, study variables, population covariance structures, study group sizes, study group means, study group variances, series correlations and missing data rates to resemble the real world situations in studies investigating human exposure to cigarette smoke. The resulting Type I error rates for testing the differences in study group means are compared. Overall, generalized estimation equations show superior Type I error control (mostly around 5% versus considerably larger than 5% for linear mixed models) in detecting the differences in study group means.

A New Class of Location Parameters and Estimators

❖Roy St Laurent, Northern Arizona University, Department of Mathematics & Statistics, Box 5717, Flagstaff, AZ 86011-5717, *roy.st.laurent@nau.edu*; Ian Harris, Southern Methodist University

Key Words: mean, median, robust estimators, influence, breakdown

We introduce a new way to relate the mean and the median. A class of parameters is defined via a functional. Each member of the class is a weighted mean of probability mass above and below the parameter. By adjusting a tuning parameter that determines the weighting, a family of location parameters is generated which includes the mean; and includes the median in a limiting sense. Applying the functional to data generates a family of location estimators. Properties of the estimators are explored and compared to those of the mean, the median and other well-known robust measures of location.

Effect of Sample Aliquot Size on the Sensitivity and Reproducibility of Clinical Assays

❖Guorong Chen, Digene Corp, 1201 Clopper Rd, Gaithersburg, MD 20878, *guorong.chen@digene.com*; Irina Nazarenko, Digene Corp

Key Words: clinical assay, probability model, sample preparation

Many available nucleic acid amplification technologies contribute to highly sensitive diagnostic assays. However, the ability to amplify fewer than 10 target copies of DNA or RNA puts new requirements for the preparation of clinical specimens. One important question is how large an aliquot is needed from a clinical specimen to reproducibly capture a certain number of target particles (for example, bacteria cells). A probability model has been developed to estimate the required aliquot size and concentration of "particles" in a clinical sample for optimal assay performance. This method can be used to calculate the probability for getting no particles in an aliquot of a given size and to determine the concentration of particles in a clinical sample to generate a reproducible sample aliquot. This model can also be applied to predict the level of variability contributed by sampling error.

Noninferiority Testing with a Given Percentage of the Control as the Noninferiority Margin

❖Paul Hshieh, Food and Drug Administration, 1401 Rockville Pike, CBER, Rockville, MD 20852, *paul.hshieh@fda.hhs.gov*

Key Words: Noninferiority, noninferiority margin, mean ratio, mean difference

In the evaluation of blood products, a given percentage (e.g., 20%) of the control mean is often used as a noninferiority (NI) margin; however, calculation of this margin using the control sample mean is problematic because doing so does not take the variability of the estimate into consideration. Furthermore, NI margin should be fixed in the design stage; it should not be computed or estimated from the current data. To resolve these problems, the noninferiority hypothesis may be formulated based on the mean ratio rather than the mean difference. Such a hypothesis may be tested using different approaches. Real examples will be discussed.

Using Spline-Enhanced ODEs for PK/PD Model Development

❖Dong Wang, University of Nebraska-Lincoln, 340 Hardin Hall, Lincoln, NE 68583-0963, *dwang3@unl.edu*; Yi Wang, University of Nebraska-Lincoln; Kent Eskridge, University of Nebraska-Lincoln; Shunpu Zhang, University of Nebraska-Lincoln

Key Words: PK/PD, spline-enhanced ODEs, penalized least squares

Motivated by using stochastic differential equations for PK/PD model development, a method for developing PK/PD models based on the spline-enhanced differential equations is proposed. The new method allows information about the appropriate model structure to be extracted directly from data, which is accomplished by inclusion of flexible nonparametric functions in the ODEs to quantify model uncertainty. Furthermore, the new method provides the basis for systematic model development by allowing time-variations in key parameters to be tracked and visualized graphically. The performance of the new method is demonstrated by means of two examples with simulated data. Compared to SDE-based method, advantages of the proposed method are: intensive computation is avoided by using penalized least squares and the model misspecification can be modified along with the improved parameter estimates.

Wavelet Analysis of Super-Frequent fMRI

❖Zibonele Valdez-Jasso, The University of Texas at Dallas, 7760 McCallum Blvd Apt 20317, Dallas, TX 75252, *zav061000@utdallas.edu*

Key Words: fMRI, Wavelets, Denoising, super-frequent fMRI

Developing a technology that is capable of obtaining brain images with high frequency is necessary for understanding neuronal events and many brain diseases. To address this issue, the Advanced Imaging Research Center at UT Southwestern pioneered a super-frequent functional magnetic resonance imaging (fMRI). This technique allows capturing changes in brain activities with 50msc time intervals and preserves the good spatial resolution of a classical FMRI. Conducted experiments revealed that the response signal level is comparable with noise level; and pronounced respiratory and cardiac signals, which contaminate the response signal. A new wavelet-based methodology has been developed and tested to urgently response to this need. Obtained results are discussed including possible decomposition of an fMRI time series into a response signal, nuisance respiratory and cardiac signals and noise.

Statistical Analysis of EEG-Sleep Patterns of Neonates

❖Alexandra Piryatinska, San Francisco State University, 1000 Continentals way apt 214, Belmont, CA 94002, *alpiryat@sfsu.edu*

Key Words: time series analysis, change point detection, cluster analysis, analysis of eeg

We analyze the EEG-sleep signal of neonates. A rigorous methodology to analyze EEG-sleep pattern are needed to assess functional brain maturation of infants. Changes in ultradian rhythm of the sleep (change between sleep stages) is one of the important characteristics related the brain maturation. We consider the EEG signal as a piecewise stationary time series. Several EEG spectral and nonlinear dynamics characteristics are being estimated. We find the characteristics that are significantly different for the different sleep stages. The nonparametric change-point detection algorithm and cluster analysis is applied to these characteristics to obtain the sleep stage separation. The groups of pre-term and full-term neonates have been compared.

155 Section on Statistics in Sports Speaker with Lunch (fee event)

Section on Statistics in Sports

Monday, July 30, 12:30 pm–1:50 pm

Is Tiger a Winner, or Is He Just Better Than Everyone?

❖Scott Berry, Berry Consultants, 3145 Chaco Canyon Drive, College Station, TX 77845, *scott@berryconsultants.com*

Key Words: Golf, Tiger Woods, Psychological Effects, Sports, Bayesian

When commentators and the sporting public talk about Tiger Woods they inevitably say he has a tremendous will to win and is mentally better than everyone else. Sports are notorious for attaching "reasons" to randomness–there has to be a reason he wins so darn much. The sports world also loves "winners." Is Tiger Woods a winner, or is he so much better than everyone else that he just wins a lot? In this talk, I use a Bayesian hierarchical model to address how good Tiger Woods is—and why he wins so much. If time permits I also will discuss what elements of his game make him so good—is it his individual measurable attributes or his intangibles?

156 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section

Monday, July 30, 12:30 pm–1:50 pm

Innovative Design and Analysis of FDA Medical Device Submissions

❖Gary Kamer, Food and Drug Administration, 1350 Piccard Drive, Rockville, MD 20850, *gary.kamer@fda.hhs.gov*

Key Words: FDA, submissions, innovative, methodologies, frequentist

This discussion will emphasize the acceptability of recently developed statistical methodologies to establish the safety and effectiveness of new medical devices seeking to be marketed in the United States. It is anticipated that there will be a broad range of methodologies discussed, along with their strengths, weaknesses, and alternatives. Concentration will be on frequentist procedures, but Bayesian issues also may be addressed.

Blocks, Strata, and Covariates: What's in a Name?

❖Veronica Taylor, Food and Drug Administration, Center for Veterinary Medicine, 7500 Standish Place HFV-105, Rockville, MD 20855, *veronica.taylor@fda.hhs.gov*

Key Words: randomized block, strata, covariate

Blocks, strata, and covariates may be used to reduce variance to provide a better test for treatment effects of interest. Sometimes the terms are used interchangeably. However, each concept has different processes and assumptions attached. We will discuss the underlying assumptions, uses in experimental design, and possible analyses when blocks, strata, or covariates are used.

The Use of Meta-Analysis in Drug Development

❖Tsushung Hua, Novartis Pharmaceuticals, One Health Plaza, Build 419 Room 2112, East Hanover, NJ 07936, *tsushung.hua@novartis.com*

Key Words: Evidence synthesis, clinical trials, combination of information

Meta-analysis (MA) is a statistical method to obtain an overall drug (treatment) effect or results from independent, but related, trials. Lately, medical journals have published more papers from meta-analyses to show the efficacy and safety of a drug or class of drugs. In this roundtable, we will discuss the method and application of MA in drug development.

Evaluating Probability of Success for Internal Decisionmaking in Early Drug Development

❖Narinder Nangia, Abbott Laboratories, 100 Abbott Park Road, Dept. R436, BLDG. AP9A-2, Abbott Park, IL 60064-6124, *narinder.nangia@abbott.com*

Key Words: Bayesian, Clinical Trials, Early Drug Development, Probability of Success, Inflammation, Oncology

There is a need to look beyond p-values in drawing meaningful inference from the analysis of phase I-II (learning stage) clinical trials data. This discussion will be on using the Bayesian approach to exploit totality of accumulated data/knowledge in a formal way for internal decisionmaking in early drug development. Examples from inflammation and oncology therapeutic areas will be considered for evaluation of probability of success for drug candidates in meeting the target product profile.

157 Business and Economics Statistics Section Roundtable with Lunch (fee event)

Business and Economics Statistics Section

Monday, July 30, 12:30 pm–1:50 pm

Statistics Teaching: Bayesian, Frequentist, United

❖Emanuel Parzen, Texas A&M University, Statistics Dept, College Station, TX 77843-3141, *eparzen@tamu.edu*

Key Words: statistical education, quantiles, posterior quantiles, confidence quantiles, united statistics, data modeling

All statisticians have problems understanding distinction between confidence intervals and credible intervals (which I call statistical inference without priors and with priors, rather than frequentist and Bayesian inference). In introductory statistics education, question is how to teach Bayesian prior methods while teaching with equal emphasis frequentist nonprior methods. We propose that the practice of statistical inference requires juggling several distributions simultaneously and a village of quantile functions of distributions. One important consequence of parallel reasoning between posterior quantiles and confidence quantiles (endpoints of confidence intervals): These provide for hypothesis testing both Bayesian and frequentist solutions; we recommend using and comparing both!

158 Section on Bayesian Statistical Science Roundtable with Lunch (fee event)

Section on Bayesian Statistical Science

Monday, July 30, 12:30 pm–1:50 pm

Dealing with Review Efficiency: A Practical Bayesian Approach

❖Xiao-li Meng, Harvard University, 1 Oxford Street 7th Floor, Dept of Statistics, Cambridge, MA 02138, *meng@stat.harvard.edu*

Key Words: Referee Process, Statistica Sinica, Publication, Promotion, Frustration, Author

The slow manuscript review process has frustrated generations of statisticians, Bayesian or not. Many have put in tremendous effort to improve review efficiency, but more collective effort is needed to fundamentally shift our slow review culture. Some believe Bayesians have solutions to every problem, especially in improving efficiency. A Bayesianly justifiable approach was recently implemented by Statistica Sinica, with a result that appears to support that belief. The median for Statistica Sinica's first review period (from the date of submission to first editorial decision sent) is now about five weeks, with 95% percentile about four months and 100% percentile about six months. Undoubtedly, this right tail is still too long, so please join me to brainstorm about ways to further improve, especially if you believe you can do better than Bayesians.

159 Section on Government Statistics Roundtable with Lunch (fee event)

Section on Government Statistics
Monday, July 30, 12:30 pm–1:50 pm

Design and Use of the IPUMS-International Data Series

❖Matthew Sobek, The University of Minnesota, 3739 47th Ave South, Minneapolis, MN 55406, *sobek@pop.umn.edu*

Key Words: census, population, database, large datasets, dissemination

The Integrated Public Use Microdata Series (IPUMS) project is compiling the world's largest public-use census database. The data system currently includes 63 census samples from 20 countries, containing the individual-level records of 185 million persons. Samples from another 25 countries will be added over the next three years. The project harmonizes the data and documentation across countries, and the web system produces data extracts that users download for analysis—all completely free of charge. This roundtable will discuss the design of the IPUMS, including key decisions of the research team and constraints imposed by international statistical offices. We also will discuss the strengths and limitations of the data and directions for future development of the series.

160 Section on Health Policy Statistics Roundtable with Lunch (fee event)

Section on Health Policy Statistics
Monday, July 30, 12:30 pm–1:50 pm

Estimation of Causal Effects in Health Services and Outcome Research Using Observational Data

❖Lirong Zhao, University of Maryland, 220 Arch Street, 12th Floor, Baltimore, MD 21201, *lzhao@rx.umaryland.edu*

Key Words: propensity score, instrumental variable

Sample selection bias is one of the major concerns in the estimation of causal effects using observational data in health services and outcome research when experimental designs are not feasible. Statisticians, epidemiologists, and econometricians have proposed many methods to adjust this bias over past decades. Propensity score (PS) and instrumental variable (IV) are two of the major methods that reduce the observable and unobservable biases, respectively. In this roundtable discussion, you will have the chance to share experiences of applications of these two methods in health policy statistics.

161 Section on Physical and Engineering Sciences Roundtables with Lunch (fee event)

Section on Physical and Engineering Sciences
Monday, July 30, 12:30 pm–1:50 pm

Sensor-Based Field Reliability Data Analysis: Opportunities and Challenges

❖Necip Doganaksoy, General Electric, GE GRC, 1 Research Circle, Niskayuna, NY 12309, *doganaksoy@research.ge.com*

Key Words: Product reliability, Automated monitoring, Multivariate modeling

The advances in sensor technologies have enabled manufacturers to track their products remotely and gather useful information to assess their health. The resulting data provides significant opportunities, as well as challenges, to statisticians in automated monitoring to identify impending failures. This is in sharp contrast to the recent past when field reliability data were mainly comprised of after-the-fact failure information. Owing to large size and high dimension of sensor-based data, data mining and multivariate statistical methods are particularly well-suited to tackle the technical challenges. In this roundtable, I will share my own experiences from recent projects involving analysis and modeling of sensor data gathered from field units. Participants are encouraged to bring their examples to stimulate a useful exchange.

Non-Negative Matrix Factorization in Statistics

❖S. Stanley Young, National Institute of Statistical Sciences, 19 T.W. Alexander Drive, P.O. Box 14006, Research Triangle Park, NC 27709-4006, *young@niss.org*

Key Words: Matrix Factorization, Non-negative Matrix Factorization, NMF, Clustering, Microarray Analysis, Text mining

Matrix factorization, singular value decomposition, is the underpinning of most linear statistical methods. Nonnegative matrix factorization, NMF, takes into account that, in many cases, two-way data tables contain only non-negative numbers. NMF is being used in analytical chemistry, image processing, text classification, blind signal separation, etc., but NMF is not widely considered in the statistics literature. There appears to be a wide scope for research on the use of NMF for statistical problems. The goal of this roundtable is to briefly present aspects of NMF and hear what is going on in this area. Participants are encouraged to be ready to discuss their own possible applications of NMF and/or analytic alternatives they have tried.

162 Section on Quality and Productivity Roundtable with Lunch (fee event)

Section on Quality and Productivity
Monday, July 30, 12:30 pm–1:50 pm

Exposing Students to Six Sigma as Part of Their University Training

❖Angie Patterson, GE Global Research, 417 Ridgeview Dr, Blacksburg, VA 24060, *patterso@research.ge.com*

Key Words: Six Sigma, Teaching, Statistics, Education, Problem Solving, University

University students can greatly benefit from exposure to Six Sigma. Statistics majors (graduate and undergraduate) will gain nonstatistical skills (leadership, teamwork, structured problemsolving) that are not emphasized in their core coursework. Majors from other disciplines will be better motivated to learn statistical methodology when taught in an applied problemsolving framework. The discussion leader will share her experience in developing and teaching a master's level course in the Virginia Tech Department of Statistics (course overview, project selection and student feedback), as well as share teaching models at other universities. This roundtable discussion will benefit those who are interested in developing a course for their university and those wishing to network and share best practices from their experience.

163 Section on Risk Analysis Roundtable with Lunch (fee event)

Section on Risk Analysis

Monday, July 30, 12:30 pm–1:50 pm

Harnessing the Power of Information in Quantitative Risk Assessment

❖Walter W. Piegorsch, The University of Arizona, BIO5 Institute, Tucson, AZ 85721, *piegorsch@math.arizona.edu*

Key Words: quantitative risk analysis, environmental risk assessment, benchmark dose, uncertainty analysis, sensitivity analysis, data mining

A broad discussion will be held on how modern information resources and technologies can be brought to bear on problems in quantitative risk analysis. Possible topics of pertinence will include potency estimation for dose-response data, excess risk estimation, benchmark analysis and benchmark dose (BMD) estimation, large-database mining, uncertainty analysis, and sensitivity analysis. The discussion will remain as flexible as possible to include the many, wide-ranging aspects of modern quantitative risk assessment.

164 Section on Statistical Consulting Roundtables with Lunch (fee event)

Section on Statistical Consulting

Monday, July 30, 12:30 pm–1:50 pm

Right Answer, Wrong Question: Characterization and Reduction of the Risk of Type III Error in Statistical Applications

❖John L. Eltinge, Bureau of Labor Statistics, OSMR, PSB 1950, 2 Massachusetts Avenue NE, Washington, DC 20212, *Eltinge.John@bls.gov*

Key Words: Adoption and diffusion of technology, Communication, Feedback loops, Opportunity cost, Project management, Risk management

2007 marks the 50th anniversary of the publication of an article by A. W. Kimball on Type III error, defined as "the error committed by giving the right answer to the wrong problem." This open roundtable will explore practical approaches to Type III error, with special emphasis on four questions: What do we mean by "wrong problem"? What are some especially notable examples of Type III error that the roundtable participants have encountered in their consulting work? In light of the abovementioned examples, what systematic approaches can statisticians take to identify potential sources of Type III error, to detect these errors quickly when they occur, and to limit the impact of these errors? What practical implications do these systematic approaches have for the training and management of statistical consultants?

165 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education

Monday, July 30, 12:30 pm–1:50 pm

Spanning the Globe: Using Data To Introduce Students to the Rest of the World

❖Robert Carver, Stonehill College, 39 Woodland Street, Sharon, MA 02067, *robert.carver@comcast.net*

Key Words: Real data, Globalization, Introductory Statistics, Data analysis

In the introductory statistics course, we have the opportunity to expand the world view of our students. The simple idea is that there are numerous good international sources of data accessible online, and there are still precious few non-USA examples in most textbooks. At this roundtable, participants will swap sites, find out about easy and not-so-easy interfaces, and compare notes about the effective use of such data in the classroom and assignments.

TA Training for Introductory Statistics Courses

❖Jessica Utts, University of California, Davis, Dept of Statistics, One Shields Ave, Davis, CA 95616, *jmutts@ucdavis.edu*

Key Words: Teaching Assistant, Statistics Education, Introductory Statistics, TA Training

The statistics education community has made great progress in providing resources and advice for teaching introductory statistics. However, one major component of most introductory courses that has not received attention is the training of teaching assistants for these courses. At this luncheon, we will share information about what is done at our institutions and discuss ideas about what we can do better.

Learning to Teach Statistics: Challenges and Suggestions

❖Candace Berrett, The Ohio State University, 3504 Prestwick Ct, Columbus, OH 43220, *cberrett@stat.ohio-state.edu*; Jacqueline Miller, The Ohio State University

Key Words: Teaching statistics, TA training

As students in statistics, most of us are expected to teach in some fashion, but we all begin with various backgrounds in teaching and in our knowledge of statistics, making a cookie cutter training program difficult. At this roundtable, we will discuss various TA training strategies, what training you received (if any), the pros and cons of your training, and other preparation that would have been beneficial. Come and enjoy lunch with the company of other TAs, find out what others are doing to prepare to teach, and get ideas to improve your own teaching.

166 Section on Statistical Graphics Roundtable with Lunch (fee event)

Section on Statistical Graphics

Monday, July 30, 12:30 pm–1:50 pm

Graphical Data Mining of Network Data

❖George Michailidis, The University of Michigan, 439 West Hall, 1085 S. University, Ann Arbor, MI 48109, *gmichail@umich.edu*

Key Words: network data, data mining, visualization, graph drawing, algorithms

Network data have become ubiquitous in many scientific fields. They are capable of representing relationships between various entities, such as people in social or epidemiological contact networks, machines in computer and communications networks, genes and proteins in biological networks, etc. They also pose various challenges due to their topological structure and large scale. Visualization is essential for exploring and analyzing such data. We will discuss techniques based on efficient graph-drawing algorithms for visualizing such data and for discovering interesting patterns and trends. They will be illustrated on a number of real-world examples from diverse fields.

167 Section on Statisticians in Defense and National Security Roundtable with Lunch (fee event)

Section on Statisticians in Defense and National Security

Monday, July 30, 12:30 pm–1:50 pm

Increasing Understanding of the Need for Statistics in Defense and Security

❖Alyson G. Wilson, Los Alamos National Laboratory, P.O. Box 1663, MS F600, Los Alamos, NM 87545, *agw@lanl.gov*

Key Words: defense, security

In 2006, the ASA president's Task Force on Statistics in Defense and Security was charged to "provide advice and a plan on how the statistical community can better serve the needs of the nation in the areas of national defense and security and identify ways in which those communities can have greater access to statistical expertise and the statistical community can learn about areas of concern in defense and security." In this roundtable, the task force chair (Alyson Wilson) and a member (Wendy Martinez) will discuss its recommendations, and ask: 1. Who are the main players in defense and security, at the policy, funding, and applications levels? 2. What are the important problems that need statistics? 3. How do we communicate statistics to the stakeholders? 4. What are the lessons learned from previous attempts to do this?

168 Section on Statistics in Epidemiology Roundtables with Lunch (fee event)

Section on Statistics in Epidemiology

Monday, July 30, 12:30 pm–1:50 pm

Gerontologic Biostatistics and the Challenges of Conducting Biomedical Research with Older Populations

❖Peter H. Van Ness, Yale University, Program on Aging, School of Medicine, 1 Church Street / 7th Fllor, New Haven, CT 06510, *peter.vanness@yale.edu*; ❖Heather G. Allore, Yale University

Key Words: Aging, Longitudinal, Missing Data, Qualitative, Design, Medicine

Multifactorial etiologies of geriatric health syndromes and multiple morbidities of older patients present challenges for designing and analyzing research studies in which they participate. The objective of this roundtable is to discuss a set of statistical issues that, while not unique to gerontologic research, constitutes a distinctive ensemble worthy of deliberate and systematic attention (e.g., design of multicomponent intervention trials, regression modeling with missing data, evaluation of clinical tradeoffs, analysis of complex longitudinal datasets, and integration of qualitative studies with quantitative analyses). The roundtable will include a discussion of the invited session devoted to the topic of gerontologic biostatistics at this year's Joint Statistical Meetings and planning for future activities for statisticians interested in this topic.

169 Section on Survey Research Methods Roundtable with Lunch (fee event)

Section on Survey Research Methods

Monday, July 30, 12:30 pm–1:50 pm

What Happened in Florida-13's Election Last Year, and What Can We Do To Improve Electoral Integrity?

❖Arlene Ash, Boston University, 720 Harrison Avenue, School of Medicine, Suite 1108, Boston, MA 02115, *aash@bu.edu*

Key Words: voting, politics, public policy, electoral integrity

The statistical evidence that poor electronic ballot design altered the outcome of the election for U.S. House of Representatives in Florida's 13th Congressional district last year is compelling. The problem (that the screen needed to be scrolled to see how to vote in this race) led to >15,000 under votes in the one county that favored the candidate who apparently lost by < 400 votes; voters in other counties faced a screen with no scrolling. Although this problem could have occurred innocently, ballot designers will likely continue to make such errors, innocently or not. We will discuss evidence from this election, reasons for the persistent failure to eliminate errors in U.S. elections (no CQI), the disconnect between statistical knowledge and political/judicial realities, and potential roles for statisticians (and the ASA) in helping to improve electoral processes and outcomes.

170 Social Statistics Section Roundtable with Lunch (fee event)

Social Statistics Section

Monday, July 30, 12:30 pm–1:50 pm

Linking Survey Data and Administrative Data for Policy Research

❖Michael Davern, The University of Minnesota, 2221 University Ave SE, Suite 345, Minneapolis, MN 55414, *daver004@umn.edu*

Key Words: survey data, administrative data, policy research, linked data

Linked survey and administrative data files have great potential for policy research, as they can mitigate known problems with the survey data, itself. By linking the survey data with administrative data, two major limitations are overcome. Survey data contain a significant amount of measurement error especially when it comes to measuring public program enrollment. Administrative data do not include data on those eligible, but not enrolled, in the program or those who would potentially be eligible, if the program rules changed. By linking the files, the eligible, but unenrolled, information can be provided by the survey data and the program enrollment data can be provided by the administrative data (correcting the survey measurement error). This work is crucial for program simulation modeling used to score federal legislation and to evaluate particular public programs.

171 Late-Breaking Session 1: Statistical Analysis of an Archeological Find

The ASA, ENAR, IMS, WNAR, SSC

Monday, July 30, 2:00 pm–3:50 pm

Statistical Analysis of an Archeological Find

❖Andrey Feuerverger, University of Toronto, Department of Statistics, Toronto, ON M6C 1W3, *andrey@utstat.toronto.edu*

An interesting data set originating from an archaeological find and the statistical problems associated with it will be described. In 1980 a burial tomb was unearthed in Jerusalem containing ossuaries (limestone coffins) bearing such inscriptions as Yeshua son of Yehosef, Marya, Yoseh—names which match those of New Testament figures, but which were otherwise in common use. We will discuss some of the statistical issues that arise in authenticating or repudiating links between this find and the New Testament family. An approach is proposed—based on `relevance' and `rareness' of names—for assessing the `surprisingness' of the observed outcome. The important role played by historical assumptions will be discussed. Some unresolved statistical questions (both applied and theoretical in nature) encountered during this work will be described.

172 Statistical Issues in High-Dimensional Omics Data and Biomarker Discovery ● ✪

Biopharmaceutical Section, ENAR, Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Systems-Based Analysis of Metabolic Profiles

❖Jacqueline M. Hughes-Oliver, North Carolina State University, Department of Statistics, Room 203C Patterson Hall, 2501 Founders Drive, Raleigh, NC 27695-8203, *hughesol@stat.ncsu.edu*

Key Words: disease diagnosis, false discovery rate, Fisher's exact test, KEGG, metabolomics, pathway

Metabolomics is emerging as an attractive component of the extensive body of platforms for systems biology. The relatively small number of metabolites and their mostly known network of interconnectivity and relation to disease make metabolomics a prime candidate for improving both disease diagnosis and treatment. Unfortunately, analysis of metabolomic data typically ignores knowledge of disease pathways, and hence the opportunity for taking full advantage of domain knowledge is lost. This work develops a simple systemized view of metabolites and their existence in pathways, and then incorporates this knowledge in analyzing the impact of pathway measures on disease status. Our focus is more interpretable diagnosis of disease occurrence.

Toward Superior Classifications with ABC Dissimilarities: A Microarray Case Study

❖Dhammika Amaratunga, Johnson & Johnson PRD, G-021, 1000 Rt. 202-S., Raritan, NJ , *damaratu@prdus.jnj.com*; Javier Cabrera, Rutgers University; Vladimir Kovtun, Rutgers University

Key Words: clustering, classification, dissimilarity, microarrays

Microarray (and other high throughput omics) data are characterized as having an incredibly large number of predictor variables, only a few of which are truly informative and are correlated in groups, and a quite small number of samples. Classification methods, when used in conjunction with a conventional dissimilarity measure, tend to have terribly low efficiencies when applied to such data. We propose a novel dissimilarity measure that overcomes this problem. The measure is obtained via an ensemble scheme that involves a series of resample classifications, where the resampling involves both predictors and samples and is weighted towards informative features of the predictor space. We demonstrate, using actual microarray data, that this produces far superior classifications and visualizations of the data.

Sample Size Estimation for Omics Experiments

❖Gregory R. Warnes, University of Rochester, 601 Elmwood Dr, Box 630, Biostatistics and Computational Biology, Rochester, NY 14534, *warnes@bst.rochester.edu*; Peng Liu, Iowa State University

Key Words: sample size, microarray, omics, experimental design, proteomics, software

"Omics" technology is widely applied in biomedical and pharmaceutical research. The huge number of individual measurements for each sample make it difficult to apply traditional sample size calculation techniques and has left most practitioners to rely on rule-of-thumb techniques. In this paper, we describe and demonstrate a simple method for performing and visualizing sample size calculations for ``omics" experiments. We then summarize simulation results showing that this method performs well, even under a variety of departures from our simplifying assumptions, provided that the underlying formula for computing sample size has been appropriately selected. Although we demonstrate sample size calculation only for the two-sample pooled t-test, it is trivial to substitute an alternative sample size formula appropriate to the problem at hand.

173 Methods to Identify and Adjust for Selection Bias in Observational Studies ● ✪

Section on Statistics in Epidemiology, Section on Health Policy Statistics, Section on Risk Analysis, Section on Teaching Statistics in the Health Sciences, WNAR

Monday, July 30, 2:00 pm–3:50 pm

Propensity Score Calibration with Validation Data

❖ Til Stürmer, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA 02120, *til.sturmer@post.harvard.edu*; Sebastian Schneeweiss, Brigham and Women's Hospital; Kenneth J. Rothman, Brigham and Women's Hospital; Jerry Avorn, Brigham and Women's Hospital; Robert Glynn, Brigham and Women's Hospital

Key Words: bias, confounding, research design, propensity score, regression calibration, validation study

Propensity scores (PS) can adjust for selection bias caused by observed risk factors for the outcome. PS calibration (PSC) can adjust for selection bias by additional risk factors based on the joint distribution of these risk factors in a validation study. We assessed the performance of PSC simulating large cohorts in which information on a risk factor is only available for a small proportion of observations. When the PS estimated based on main study risk factors (error-prone) is independent of the outcome given the PS including additional risk factors from the validation study (gold standard), PSC leads to a bias reduction between 73% and 110%. Otherwise, PSC can lead to an increase of bias. PSC is valid when the direction of confounding due to selection bias by the additional risk factors in the validation study is the same as the one based on risk factors available in the main study.

Bayesian Methods to Acknowledge Confounding and Mismeasurement

❖ Paul Gustafson, University of British Columbia, Dept of Statistics, 333-6356 Agricultural Rd., Vancouver, BC V6T1Z2 Canada, *gustaf@stat.ubc.ca*; Lawrence McCandless, University of British Columbia

Key Words: Bayesian methods, Observational studies, Confounding, Selection bias

In observational studies, there are many ways in which the data we actually have can differ from the data we wish we had. For instance, relevant variables may be poorly measured, or even entirely absent. It may be possible to construct a reasonable model for how the actual data relate to the ideal data, as an adjunct to the model of interest for the ideal data. Often, however, the implied model for the actual data is not identified. We discuss the mechanics of Bayesian inference when the model is nonidentified, and the added value of a Bayesian analysis relative to a sensitivity analysis in this context. Also, in the specific case of unobserved confounding, we consider the impact of an a priori judgment that observed and unobserved confounders are exchangeable.

Instrumental Variables in Pharmacoepidemiology

❖ M. Alan Brookhart, Brigham and Women's Hospital, 1620 Tremont St, Harvard Medical School, Suite 3030, Boston, MA 02138, *abrookhart@rics.bwh.harvard.edu*

Key Words: pharmacoepidemiology, instrumental variables

Post-marketing studies of the comparative safety and effectiveness of prescription drugs are often based on health care claims data files. These files contain drug dispensing and clinical outcome information on entire populations followed over extended periods of time. Because these data were collected for administrative purposes rather than for research, they do not capture many important variables that physicians use to make prescribing decisions. Unmeasured confounding, therefore, is particularly worrisome in observational studies of drugs in health care claims data. Statistical methods based on instrumental variables present one possible approach to this problem, provided that suitable instruments can be identified. I describe several potential instruments for use in observational studies of drugs and discuss some of their limitations.

174 Advances in Contemporary Nonparametric Methods ✪

Section on Nonparametric Statistics

Monday, July 30, 2:00 pm–3:50 pm

On Least-Squares Fitting for Spatial Point Processes

❖ Michael Sherman, Texas A&M University, Dept. of Statistics, Blocker Building, College station, TX 77843-3143, *sherman@stat.tamu.edu*; Yongtao Guan, Yale University

Key Words: K-function, least squares estimator, subsampling, spatial point process

The K-function is a popular tool to fit spatial point process models due to its simplicity and applicability to a wide variety of point process models. In this work we study the properties of least squares estimators of model parameters and propose a new method of model fitting via the K-function using subsampling. We demonstrate consistency and asymptotic normality of our estimators of model parameters and compare the efficiency of our procedure with existing ones. This is done through asymptotic theory, simulation experiments and an application to a data set on long leaf pine trees.

A GCV Approach for Bandwidth Selection in PET Image Reconstruction

❖ Ranjan Maitra, Iowa State University, Department of Statistics, 123 Snedecor Hall, Ames, IA 50011-1210, *maitra@iastate.edu*

Key Words: generalized deconvolution, smoothed least squares, filtered backprojection, circulant matrix, Fourier Transforms, BCCB matrices

The problem of bandwidth estimation for filtered backprojection (FBP) in Positron Emission Tomography (PET) image reconstruction is studied in the context of generalized deconvolution. A generalized cross-validation (GCV) approach is presented. Results on eigendecomposition of symmetric one-dimensional and two-dimensional circulant matrices are derived and computational implications of the above in making reconstruction bandwidth estimation a simple extension of standard FBP are developed. The methodology is studied using numerical simulations on a class of idealized one-dimensional generalized deconvolution problems and PET phantom data. In each case, performance evaluations are excellent and in real-time, proving the potential viability of this approach in a practical setting. This research is supported in part by (NSF) CAREER DMS #0437555 and the NIH DC-006740.

Minimax Estimators of the Coverage Probability of the Impermissible Error for a Location Family

❖ Miguel A. Arcones, Binghamton University, Department of Mathematical Sciences, Binghamton, NY 13902, *arcones@math.binghamton.edu*

Key Words: Error of estimation, impermissible error, confidence interval, location family, minimax, translation equivariant

We consider estimation for a multivariate location family. Between all confidence regions with volume less than a fixed value L, we find the equivariant confidence region with the biggest coverage probability. This region maximizes the infimum of the coverage probability over all confidence regions with volume less than L. As an application, we find an estimator of parameter of location with the property that minimizes the supremum of the probability that the error of the estimation exceeds a fixed constant. We also find a confidence region and an estimator having the previous properties, but based on the m.l.e. We find the Bahadur slope of the two obtained estimators. We show that except for certain families of distributions, the estimator based on the whole sample is superior to the estimator based on the m.l.e. Hence, we get that m.l.e.'s are not asymptotically sufficient.

Higher-Order Properties of Block Bootstrap Confidence Intervals

❖Soumendra N. Lahiri, Texas A&M University, 405D Blocker, 3143, College Station, TX 77843, *snlahiri@stat.tamu.edu*

Key Words: percentile-t, coverage probability, stationarity, studentization, weak dependence

Higher-order properties of one- and two-sided block bootstrap confidence intervals based on weakly dependent time series data will be presented. A comparison of the results with the bootstrap confidence intervals under independence will be given. In particular, effects of the studentizing factor on the accuracy of the percentile-t CIs in the dependent case will be discussed. Results from a simulation study on finite sample properties of the block bootstrap confidence intervals will also be presented.

175 The Transition from Undergraduate to Graduate Study of Statistics ✪

Mu Sigma Rho, Section on Statistical Education

Monday, July 30, 2:00 pm–3:50 pm

Inspiring and Challenging Students in the Mathematical Sciences

❖William Y. Velez, The University of Arizona, Department of Mathematics, 617 N Santa Rita, Tucson, AZ 85721, *velez@math.arizona.edu*

Key Words: mentoring, advising, recruiting

Increasing the number of students pursuing advanced mathematical/statistical studies is an important issue that must be addressed by mathematics/statistics departments. How do we convince students that further mathematical studies are to their benefit? Where do we begin to find these students? As opposed to the other sciences, mathematics has a pool of students that arrive every year, the calculus students. Hundreds of thousands of students take calculus each year, yet few of these students decide to pursue further mathematical studies. If departments could convince more calculus students to continue on in mathematics, we would greatly increase the pool of students capable of pursuing graduate studies in mathematics/statistics. This talk will describe the efforts of the mathematics department at the University of Arizona that resulted in doubling the number of mathematics majors.

Undergraduate Preparation

❖L. Marlin Eby, Messiah College, P.O. Box 3041, One College Avenue, Grantham, PA 17027-9800, *eby@messiah.edu*

Key Words: mentoring, courses, graduate potential, graduate programs

Successful preparation usually begins early in the undergraduate program. Topics to be discussed include creating student interest in Statistics as a discipline and a career, identifying students with graduate school potential, the role of mentoring, relevant courses and internships, guiding students in choosing graduate programs, and followup with students after they enter graduate school. The author will draw on his successful experiences as a statistician within the mathematical sciences department at a college of the liberal and applied arts and sciences.

The First-Year Experience

❖Matthew Ritter, North Carolina State University, Department of Statistics, Box 8203, Raleigh, NC 27695-8203, *mbritter@ncsu.edu*

Key Words: graduate education, transition points, mentoring, retention, advising

A successful student's perspective on the transition from undergraduate to graduate studies in statistics will be presented. The transition involves changes in level of academic difficulty, cultural changes in how students interact and study and workload issues including fellowships or assistantships. Strategies that worked or did not work will be discussed including peer mentoring, academic advising, "summer camp" and bridge courses. We will also discuss how the smoothness or difficulty of this transition affects student engagement and retention further down the line.

The First-Year of Graduate Studies in Statistics

❖Pam Arroway, North Carolina State University, CB 8203, Raleigh, NC 27695, *pam_arroway@ncsu.edu*

Key Words: graduate school, transition

Many US students find the first year of graduate studies in Statistics to be a big leap from finishing undergraduate studies. The workload, the expectations of the professors and stiff competition from international students are not what they are expecting. In this talk, we discuss strategies for aiding students in making this transition. A smoother transition in this first year of a Master's may also help lay ground work for retaining students for doctoral study.

176 State of the Art in Mixed Models Analysis ● ✪

ENAR, Section on Statistics in Epidemiology, Biometrics Section, Biopharmaceutical Section, WNAR

Monday, July 30, 2:00 pm–3:50 pm

Sample Size and Power Calculations in Late-Phase Clinical Trials with a Longitudinal Endpoint

❖Jose Pinheiro, Novartis Pharmaceuticals, One Health Plaza, East Hanover, NJ 07928, *jose.pinheiro@novartis.com*

Key Words: Correlated data, Covariance structure, Drop-outs, LME models, mixed-effects models, sensitivity analysis

Statistical analyses for primary endpoints in confirmatory clinical trials need to be clearly specified in the study protocol, before any data is collected. This includes, of course, power and sample size calculations for the specified methodology. Although this is mostly not an issue when simple methods are used, it can lead to difficulties when more complex approaches are considered. We will discuss issues involved in sample size and power calculations when a linear mixed-effects model is used to analyze a longitudinal clinical endpoint. These include, among others, considerations about drop-outs, uncertainty about underlying covariance structures, and sensitivity to misspecification of prior values for the model parameters. A real clinical trial on an neurological disease indication will be used to motivate and illustrate the methods presented.

Internal Pilot Designs and Mixed Models

❖Matthew Gurka, University of Virginia, P.O. Box 800717, Charlottesville, VA 22908-0717, *mgurka@virginia.edu*; Christopher S. Coffey, The University of Alabama at Birmingham; Keith E. Muller, University of Florida

Key Words: Internal pilot, mixed model, repeated measures, sample-size re-estimation, adaptive designs, power

An internal pilot (IP) design uses interim sample size analysis to adjust the final number of observations. Although most research on IP designs has been conducted for univariate models, the common use of multivariate data naturally motivates interest in extending IP designs to more complex models. Mixed models are a popular tool for multivariate Gaussian data exhibiting missing values or mistimed observations. However, the development of IP methods for mixed models requires fixed sample tests that guarantee accuracy of inference in small samples and accurate power algorithms. Initially, we describe how to extend exact univariate IP methods to a restricted class of linear mixed models. We describe an example in medical imaging that directly benefits from this extension. We then discuss ongoing and future work for extending the use of such methods to more general mixed model settings.

Assessing the Performance of a Symmetric Divergence Information Criterion for Selecting the Best Linear Mixed Model

❖Lloyd J. Edwards, The University of North Carolina at Chapel Hill, Dept. of Biostatistics, 3105H McGavran-Greenberg, CB# 7420, Chapel Hill, NC 27599-7420, *Lloyd_Edwards@unc.edu*; Anita Abraham, The University of North Carolina at Chapel Hill

Key Words: Information Criterion, Directed Divergence, Symmetric Divergence, Model Selection, Mixed Model

The AIC and BIC are the most popular information criteria for model selection in the linear mixed model. Both criteria are directed divergences, i.e., they are asymmetric divergences for discriminating from observations on a true model in favor of an approximating model. Reversing the roles of the true model and approximating model provides an alternate directed divergence. The sum of the two directed divergences forms Kullback's symmetric divergence. Using simulation studies, we assess the performance of a form of Kullback's symmetric divergence proposed by Cavanaugh (1999), KIC, for selecting the best linear mixed model. We consider the KIC for selecting the best mean model and separately the best covariance model in the linear mixed model.

177 Machine Learning and Optimization

IMS, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences
Monday, July 30, 2:00 pm–3:50 pm

Support Vector Machines for Structured Outputs

❖Thorsten Joachims, Cornell University, 4153 Upson Hall, Ithaca, NY 14853, *tj@cs.cornell.edu*

Key Words: Support Vector Machines, Multivariate Regression, Machine Learning

Over the last decade, newly developed machine learning methods like Boosting and Support Vector Machine (SVM) have focused on univariate classification and regression. Can these results be transferred to multivariate prediction problems, where the goal is to predict complex objects like trees, sequences, or orderings? Such problems arise, for example, when a natural language parser needs to predict the parse tree for a sentence, when a navigation assistant needs to predict a route, or when a search engine needs to predict a ranking. This talk will discuss an SVM approach to predicting complex objects. It generalizes the idea of margins to complex prediction problems and a large range of loss functions. While the resulting training problems have exponential size, there is a simple algorithm that allows training in polynomial time. Empirical results will be given for several examples.

Large-Scale Covariance Selection by Chordal Embedding

❖Lieven Vandenberghe, University of California, Los Angeles, Electrical Engineering Department, 68-119 Engineering IV, Los Angeles, CA 90095-1594, *vandenbe@ee.ucla.edu*; Joachim Dahl, Aalborg University

Key Words: covariance selection, convex optimization, graphical models, chordal graphs, semidefinite programming

We discuss algorithms for maximum likelihood estimation of Gaussian graphical models with conditional independence constraints, also known as covariance selection. The problem can be formulated as an unconstrained convex optimization problem and has a well-known closed-form solution if the underlying graph is chordal. Our focus will be on convex optimization algorithms for problems with non-chordal graphs. We first derive efficient methods for evaluating and inverting the gradient and Hessian of the log-likelihood function when the underlying graph is chordal. The algorithms are formulated as simple recursions on a clique tree. We then use these results to obtain efficient implementations of Newton's method and the conjugate gradient method for large non-chordal graphs, by embedding the graph in a chordal graph. We also discuss connections with sparse semidefinite programming.

Model Selection for SVMs through Bilevel Optimization

❖Gautam Kunapuli, Rensselaer Polytechnic Institute, Department of Mathematical Sciences, 110 8th Street, Troy, NY 12180, *kunapg@rpi.edu*

While support vector and kernel methods have been successfully applied, there are still open questions regarding model selection. Here, we consider the problem of model selection for support vector classification and regression through cross-validation formulated as a bilevel program. This allows for versatility in handling several machine-learning goals simultaneously and efficiently: optimal choice of model parameters, feature selection for dimensionality reduction, and kernelization to handle nonlinear datasets. The bilevel formulation is converted to a mathematical program with equilibrium constraints (MPECs). Unlike SVMs, which are convex problems, the resulting problem is nonconvex. We discuss how these challenges can be solved using decomposition and successive linearization and present results.

178 Analyzing Multiple-Response Categorical Data from Complex Survey Designs ● ✪

Section on Survey Research Methods, SSC
Monday, July 30, 2:00 pm–3:50 pm

Simple Practitioner-Oriented Hypothesis Tests for Various Types of Multiple Response Survey Data

❖D. Roland Thomas, Carleton University, Sprott School of Business, Ottawa, PA K1S 2T9 Canada, *rthomas@sprott.carleton.ca*; Yves J. Decady, Statistics Canada

Key Words: Multiple response, pick any, Rao-Scott tests

The focus of the paper will be tests of association featuring various types of multiple response data, including "pick any out of r" data in which respondents are invited to select all items that apply to them from a list of r items, "pick k out of r" data in which respondents are limited to selecting k items, along with variants for which respondents are asked to "order r" and "order k out of r" items. Also discussed will be tests where the original multiple response data have been aggregated into a smaller number of categories. Rao-Scott adjusted chi-squared tests will be used, with a view to developing simple tests that can be implemented by practitioners using standard software.

Estimation and Testing for Association with Multiple-Response Categorical Variables from Complex Surveys

❖Christopher R. Bilder, University of Nebraska-Lincoln, Department of Statistics, 340 Hardin Hall, north wing, Lincoln, NE 68583-0963, *chris@chrisbilder.com*; Thomas M. Loughin, Simon Fraser University

Key Words: choose all that apply, correlated binary data, generalized loglinear model, marginal model, pick any, simultaneous pairwise marginal independence

Many survey questions invite respondents to "choose all that apply" from a series of items. Variables that summarize the resulting data are often referred to as multiple-response categorical variables (MRCVs). Analyzing associations among MRCVs is a challenge because each one is a vector of correlated binary responses, and associations between these vectors can take on many possible forms. This talk will show how one can model these MRCVs in situations where the data has arisen from a complex survey sampling design. Generalized loglinear models are proposed for analyzing the associations between items of different MRCVs. Model comparison test statistics along with their derived asymptotic distributions are presented in order to choose a good fitting model. Model-based estimates of the odds ratios and their corresponding standard errors are provided.

Bayesian Inference for a Stratified Categorical Variable Allowing All Possible Category Choices

Balgobin Nandram, Worcester Polytechnic Institute; ❖Myron Katzoff, National Center for Health Statistics, Office of Research and Methodology, 3311 Toledo Road, Hyattsville, MD 20782, *mjk5@cdc.gov*; Ma Criselda S. Toto, Worcester Polytechnic Institute

Key Words: Bayes factor, Monte Carlo integration, Multinomial-Dirichlet model, Sparse table, Business survey, Random samples

We consider problems of inference from survey data for proportions when sample individuals have been asked to mark all responses that apply to them, the table with mutually exclusive categories is sparse, the number of individuals to whom none of the listed categories apply is missing and the category proportions are to be compared across population strata. We consider an example from the Kansas Farm Survey. We use a Bayesian product multinomial-Dirichlet model to fit the count data both within and across education levels. We estimate the proportions of individuals with each choice; show how to estimate the most frequently indicated choice; and show, using the Bayes factor, how to test that these proportions are the same over different levels of farmers' education. Our Bayesian procedure uses a sampling based method with independent samples.

179 Fifty Years of the National Health Interview Survey ●

Section on Health Policy Statistics, Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section, Section on Statistics in Epidemiology

Monday, July 30, 2:00 pm–3:50 pm

The Evolution of Concepts in the National Health Interview Survey

❖Jennifer Madans, National Center for Health Statistics, 3311 Toledo Rd, Hyattsville, MD 20782, *JMadans@cdc.gov*

This talk will describe how conceptual approaches to developing and revising the NHIS questionnaire have evolved over fifty years. For example, questions in early versions of the NHIS asked about an individual's perceived health status and its impact on functional limitations, but did not ask about diagnoses of medical conditions, whereas today's NHIS collects information on diagnoses by medical professionals. The changing approaches to collecting NHIS' respondent-reported information will be described in relation to changes in theories of public health.

Policy Uses of National Health Interview Survey (NHIS) Data on Health Insurance Coverage

❖Amy J. Davidoff, University of Maryland, School of Pharmacy, Dept of Pharmaceutical Health Services Research, 220 Arch Street, 12th floor, Baltimore, MD 21201, *adavidof@rx.umaryland.edu*

Key Words: Health insurance, National Health Interview Survey, policy analysis

Over the past decade, the NHIS has been an important source of data for estimates of insurance coverage and evaluations of policy initiatives affecting eligibility and takeup of public insurance, and offers of, and enrollment in private group plans. The NHIS permits a focus on population subgroups defined by health status, and examination of links between insurance and access to care and healthcare service use. With its commitment to provide early estimates, rapid public release of microdata, and flexibility in questionnaire design to capture new policy initiatives such as Medicare Part D, NCHS has positioned the NHIS to be a critical resource for evaluation of future policy initiatives. Examples from the literature will be used to illustrate the role played by the NHIS. Future information needs and the role of the NHIS as a key resource for researchers will be discussed.

The National Health Interview Survey in the Present and in the Future

❖Jane F. Gentleman, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *JGentleman@cdc.gov*

Key Words: National Health Interview Survey, 50th anniversary of the NHIS

The National Health Interview Survey (NHIS), conducted for 50 years by the National Center for Health Statistics, is a multi-purpose health survey of the civilian, noninstitutionalized household population of the U.S. Its stable core and annually-changing supplementary specialized questions provide rich multivariate data for numerous types of users. Core topics include health status, limitation of activities, health care access and utilization, health insurance coverage, socio-demographic information, and income. Supplement sponsors have included the Centers for Disease Control and Prevention, numerous Institutes from the National Institutes of

Health, and various other government agencies. This talk will describe the rich multivariate NHIS data, the analytic products produced from those data, the annual NHIS supplements, and expectations and hopes for the future of the NHIS.

Fifty Years of the National Health Interview Survey

❖Nancy Breen, National Cancer Institute, Health Services and Economics Branch, 6130 Executive Blvd., MSC 7344, Rockville, MD 20852-7344, *breenn@mail.nih.gov*

Key Words: cancer control, National Health Interview Survey

Since 1987, the National Cancer Institute has collaborated with NCHS to develop and administer cancer control supplements to the NHIS. Data resulting from this collaboration provide cancer screening and tobacco control estimates for measuring healthy people goals and other national benchmarks. These NHIS supplements have also enhanced the availability of data for other aspects of cancer control, including tobacco consumption, other risk factors, and family history of cancer. Dr. Breen will present some trends and how these data have been analyzed to enhance our understanding of cancer control. Examples will be taken from cancer screening, tobacco control and other cross-cutting risk factor data available in the NHIS. Using the NHIS has been especially productive to NCI because of the wide range of co-variates collected on the NHIS, the nation's primary omnibus health survey.

180 Financial Econometrics

IMS

Monday, July 30, 2:00 pm–3:50 pm

Local Parametric Inference in the Hidden Semimartingale Model

❖Per A. Mykland, The University of Chicago, Department of Statistics, 5734 University Avenue, Chicago, IL 60637, *mykland@pascal.uchicago.edu*

Key Words: local likelihood, high frequency data, financial data, semimartingale, contiguity

The availability of high frequency data for financial instruments has opened the possibility of accurately determining volatility in small time periods, such as one day. Recent work on such estimation indicates that it is necessary to analyze the data with a hidden semimartingale model, typically by the addition of measurement error. The current talk discusses likelihood approaches to inference in such models.

Parameter Estimation of Diffusion Process with Sequential Monte Carlo

❖Rong Chen, University of Illinois at Chicago, Dept of Information and Decision Sciences, (MC 294), Chicago, IL 60607, *rongchen@uic.edu*

Key Words: MLE, diffusion process, importance sampling

In financial markets and other applications, continuous time diffusion processes are often observed at discrete time. For nonlinear processes, the likelihood function of the parameters is often much easier to evaluate with continuously observed paths of the process than with discretely observed paths. In this paper we propose to use a modified version of sequential Monte Carlo method to sample the continuous path based on discretely observed observations. The sampled continuous paths are then used to estimate the likelihood function surface of the parameters. An enhanced importance sampling scheme is used to efficiently estimate the entire likelihood surface with a small number of samples. Empirical study and real applications are presented.

Heterogeneous Autoregressive Realized Volatility Model

❖Yazhen Wang, University of Connecticut, Dept. of Statistics, 215 Glenbrook Road, Ubox 4120, Storrs, CT 06269, *yzwang@stat.uconn.edu*

Key Words: High Frequency, Realized Volatility, Dynamic Model

Volatilities of asset returns are pivotal for many issues in financial economics. The availability of high-frequency intraday data should allow us to estimate volatility more accurately. Realized volatility is often used to estimate integrated volatility. To obtain better volatility estimation and forecast, some autoregressive structure of realized volatility is proposed in the literature. This talk will present my recent work on heterogeneous autoregressive models of realized volatility.

181 Large-Scale Data Mining

Section on Statistical Computing, Section on Statisticians in Defense and National Security, Section on Quality and Productivity, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Statistical Graphics

Monday, July 30, 2:00 pm–3:50 pm

A New Family of Link Functions Extending Logistic Regression

❖William DuMouchel, Lincoln Technologies, 6641 E. Circulo Otono, Tucson, AZ 85750, *william.dumouchel@lincolntechnologies.com*

Key Words: additive, multiplicative, AERS, pharmacovigilance, polytherapy

The logistic regression link function is approximately linear for predicted probabilities between 0.2 and 0.8, but not when fitted values are very small or very large, as happens when modeling the occurrence of rare events. In the latter case, it is advantageous to have a model that can fit situations both where the effects of multiple predictors are more nearly additive or more nearly multiplicative. This paper introduces an extension of logistic regression with a continuous family of sigmoid link functions, including the standard logistic as a special case, that allows selection of an arbitrary region of near linearity (fitted probability at the point of inflection). The utility of this extended family is illustrated with analyses of the adverse drug reaction database, AERS, maintained by the U.S. Food and Drug Administration.

Statistics and Search Engines

❖Daryl Pregibon, Google, 76 9th Avenue, New York, NY 10011-5225, *daryl@google.com*

Key Words: Search engines, large data sets, data mining

Search engines (SEs) are ubiquitous yet the role that statistics and data mining play in their care and feeding is not well known. We introduce several examples to illustrate the scale and scope of statistical applications used in modern search engines.

A Poor Man's View of Data Mining

❖William F. Szewczyk, National Security Agency, Mathematics Research Group, R&E Building Suite 6515, Ft. Meade, MD 20755-6515, *wfszewczyk@gmail.com*

Key Words: datamining, data analysis

Storing data and indexing is fine if you have the luxury of time and you know what you're looking for. But if you want to do exploratory data analysis of large data sets, or if what you're interested in changes over time, or the time between receiving the data and making a decision is short, then these solutions become too expensive. In this talk I will present an alternative approach that answers these challenges, synthesizing collection and processing in a unique way. This new approach requires a change in how well-known statistical methods are implemented and how data is analyzed.

182 Missing Data in Longitudinal Studies ● ✪

SSC, ENAR, Section on Statistics in Epidemiology, Biometrics Section, Biopharmaceutical Section, Section on Survey Research Methods, WNAR

Monday, July 30, 2:00 pm–3:50 pm

Mixed Nonhomogeneous Poisson Process Spline Models for the Analysis of Recurrent Event Panel Data

❖ Jason Nielsen, Simon Fraser University, Department of Statistics and Actuarial Science, 116 4232 Sardis Street, Burnaby, BC V5H 1K5 Canada, *jdn@stat.sfu.ca*; Charmaine Dean, Simon Fraser University

Key Words: Discrete Mixture, Frailty, Life History Data, Overdispersion, Random Effects, Longitudinal Data

A flexible semiparametric model for analyzing longitudinal panel count data is presented. Panel count data refers here to count data on recurrent events collected as the number of events which have occurred within specific follow-up periods. The model assumes that the counts for each subject are generated by a nonhomogeneous Poisson process with a smooth intensity function. Such smooth intensities are modeled with adaptive splines. Both random and discrete mixtures of intensities are considered to account for complex correlation structures, heterogeneity and hidden subpopulations common to this type of data. An estimating equation approach to inference requiring only low moment assumptions is developed and the method is illustrated on several datasets.

Reducing the Bias of Between- Within-Cluster Covariate Methods When Data Are Missing at Random

❖ John Neuhaus, University of California, San Francisco, 185 Berry Street, Lobby 4, Suite 5700, San Francisco, CA 94107-1762, *john@biostat.ucsf.edu*; Charles E. McCulloch, University of California, San Francisco

Key Words: Covariate decompositions, generalized linear mixed models, conditional likelihood

Generalized linear mixed models that partition covariates into between- and within-cluster components can provide effective analysis of longitudinal data in settings where covariates or responses are missing completely at random. However, like conditional likelihood methods, such between/within cluster approaches can yield inconsistent covariate effect estimates when data are missing at random. This talk describes and evaluates several strategies, including weighted methods, to reduce bias when data are missing at random. We illustrate these methods with simulation studies and fits to example data.

Semiparametric Methods for Clustered Binary Data

❖ Grace Y. Yi, University of Waterloo, Department of Statistics and Actuarial Science, Waterloo, ON N2L 3G1 Canada, *yyi@uwaterloo.ca*; Wenqing He, University of Western Ontario; Hua Liang, University of Rochester

Key Words: Binary data, Clustered data, Estimating Equation, Missing Data, Semiparametric regression

Clustered binary data arise commonly in practice, and generalized estimating equations methods are frequently used to analyze such data. The generalized linear model is often used to modulate the mean response in which covariates pertaining to the responses are present in a linear form through a link function. In practice, however, the relationship between the mean response and covariates may be very complex and linearity may not be adequate to capture that relationship. Under this circumstance, semiparametric regression with both linear and nonlinear terms included may be more flexible to facilitate the relationship between the response and covariates. In this talk, I will discuss semiparametric regression methods for analyzing clustered binary data with the interest centering on estimating both the mean and association parameters. Numerical studies will be reported.

Second-Order Least Squares Estimation for Nonlinear Mixed Effects Models

❖ Liqun Wang, University of Manitoba, Department of Statistics, Winnipeg, MB R3T 2N2 Canada, *wangl1@cc.umanitoba.ca*

Key Words: Semiparametric model, non-normal random effects, least squares method, simulation-based estimation, longitudinal data, exact consistency

The paper deals with nonlinear mixed effects models where the random effects have a general parametric distribution (not necessarily normal) and the distributions of other variables and random errors are nonparametric. I propose a second-order least squares estimator based on the first two marginal moments of the response variable. I also propose a simulation-based estimator when the closed forms of the marginal moments are not available. Both estimators are consistent and asymptotically normally distributed under fairly general conditions. Monte Carlo simulation studies show that the proposed estimators perform well for relatively small sample sizes. Compared to the likelihood approach, the new methods are computationally feasible and do not rely on the normality assumption.

183 Summer Institutes for Training in Biostatistics (SIBS): Addressing the Biostatistician Shortage

Section on Teaching Statistics in the Health Sciences, ENAR, Biometrics Section, Section on Statistical Education

Monday, July 30, 2:00 pm–3:50 pm

Summer Institutes for Training in Biostatistics (SIBS): Addressing the Biostatistician Shortage

❖ Song Yang, National Heart, Lung, and Blood Institute, NIH, 6701 Rockledge Drive, MSC 7938, Bethesda, MD 20892, *yangso@nhlbi.nih.gov*; ❖ Ronald E. Gangnon, University of Wisconsin-Madison, 207 WARF Office Building, 610 Walnut Street, Madison, WI 53726-2397, *ronald@biostat.wisc.edu*; ❖ Lisa M. Sullivan, Boston University, Department of Biostatistics, 715 Albany Street, Boston, MA 02118, *lsull@bu.edu*; ❖ Kimberly S. Weems, North Carolina State University, Department of Statistics, Box 8203, Raleigh, NC 27695-8203, *ksweems@stat.ncsu.edu*

Key Words: teaching, undergraduate, career options

There has long been a shortage of well-trained biostatisticians to collaborate with scientists in academia, industry, and government. The shortage is expected to worsen as senior biostatisticians retire. Few undergraduate students majoring in quantitative sciences are aware of the career opportunities afforded by biostatistics. In 2003 the NIH issued a request for applications for SIBS. The short-term goal of this initiative is to encourage quantitative students to attend graduate school in biostatistics. The long-term goal is to increase the numbers of practicing biostatisticians. Three SIBS were funded and trained students in 2004-2006. Many graduates of these programs have enrolled in top biostatistics programs and most intend to study biostatistics. Discussion will include the need for the SIBS program, comparisons between the programs, measures of success, and future plans.

184 Rising from the Ashes of the SIPP: Opportunities and Challenges of the New Dynamics of Economic Well-Being System

Social Statistics Section, Section on Government Statistics, Section on Survey Research Methods

Monday, July 30, 2:00 pm–3:50 pm

Rising from the Ashes of the SIPP: Opportunities and Challenges of the New Dynamics of Economic Well-Being System

❖David Johnson, U.S. Census Bureau, Washington, DC 20244, *david.s.johnson@census.gov*; ❖Constance F. Citro, National Academy of Sciences, 500 5th Street, NW, Committee on National Statistics, Washington, DC 22001, *ccitro@nas.edu*; ❖Howard Iams, Social Security Administration, 500 E Street SW Room 914, Washington, DC 20254-001, *Howard.M.Iams@ssa.gov*; ❖Martin David, University of Wisconsin-Madison, 2603 Middleton Beach Rd, Middleton, WI 53562, *david@ssc.wisc.edu*; ❖John Czajka, Mathematica Policy Research, Inc., 600 Maryland Ave. SW, Suite 550, Washington, DC 20024-2512, *jczajka@mathematica-mpr.com*

Key Words: SIPP, DEWS

The Survey of Income and Program Participation (SIPP) stopped collecting data in 2006 after having been in the field monthly since 1983. The SIPP had a long preparation phase - from 1976-1981, where it tested questionnaires, recall periods, respondent types, and processing systems. Reviewed by a National Academy of Sciences panel in 1993 and recommended as the data set to improve the measurement of income, poverty, and program participation on several occasions, SIPP, nevertheless, ended without achieving these goals. The distinguished panel, with substantial historical and practical knowledge of the SIPP program, will discuss the program goals, expectations, and disappointments, and discuss them in the context of the program designed to replace the SIPP, the Dynamics of Economic Well-being System (DEWS). The current plans and goals for the DEWS will be also described and discussed.

185 Invited Poster Presentations

Section on Statistics and the Environment, ENAR, IMS, SSC, WNAR, Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Estimation of Covariogram Models by Minimum Prediction Contrast for Kriging

❖Marc Genton, University of Geneva, Dept of Econometrics UNIGE, Bd du Pont d Arve 40, Geneva 4, CH-1211 Switzerland, *genton@stat.tamu.edu*; Olivier Perrin, University of Toulouse 1

Key Words: Contrast, Covariogram, Kriging, Prediction, Space-time, Wind speed

The method of spatial interpolation called kriging is based on the covariance structure of the data. Estimation of the parameters of a covariogram model are typically carried out by least squares fitting or by maximum likelihood, usually under normality for the latter case. We propose a new approach for the estimation of covariogram models based on the minimization of the integral of a contrast over the spatial domain of interest. The contrast is obtained from the variance of the difference between the kriging predictor based on the empirical covariogram and the one based on the parametric covariogram model. We derive properties of the resulting estimators from minimum contrast theory and perform various simulation studies to demonstrate the performances of our proposal. We apply our methodology to the estimation of space-time covariogram models for a dataset of Irish wind speed.

Experiments on the Earth: Smoothing and Spatial Statistics for Geophysical Applications Using Fields

❖Stephan Sain, National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, CO 80305, *ssain@ucar.edu*; Reinhard Furrer, Colorado School of Mines; Douglas Nychka, National Center for Atmospheric Research

Key Words: Universal kriging, Smoothing splines, Sparse matrices

The geosciences have traditionally demanded sophisticated uses of statistics especially as it applies to data that is correlated over space and time and the need to estimate functional relationships among geophysical quantities. Some recent trends in the analysis of geophysical processes include much larger data sets that may be derived from several different sources and also the need to draw complicated inferences from spatial fields. This poster addresses some of these issues through a tutorial on the R packages fields and spam and by example with data analysis problems drawn from the geosciences.

Stochastic Space-Time Regional Rainfall Modeling Adapted to Historical Rain Gauge Data

❖Zepu Zhang, The University of Chicago, 5734 S. Ellis Ave, Chicago, IL 60637, *zpzhang@stanfordalumni.org*; Paul Switzer, Stanford University

Key Words: Boolean model, precipitation, space-time structures, rain gauge data, stereology, stochastic hydrology

Stochastic rainfall models are important tools both for practical issues and in studies of weather- and climate-sensitive systems. We propose an event-based model, continuous in space (2-D) and time, that describes regional-scale, ground-observed storms by a Boolean random field of rain patches. The model creates complex space-time structures with a mathematically tractable framework. The estimation method relates temporal observations at fixed sites to the movement of the model storm rain field, thereby making historical rain gauge data suitable for model fitting. The model is estimated using hourly historical data at eight rain gauges in Alabama and tested for its capabilities in capturing statistical characteristics of the historical data, including rainfall intensity distribution and extremes, temporal correlation, effects of temporal aggregation, spatial coverage and correlation.

Black Diamond Ski Runs and the Fields R Package

❖Douglas Nychka, National Center for Atmospheric Research, Box 3000, Boulder, CO 80307-3000, *nychka@ucar.edu*

Key Words: spatial statistics, Kriging, thin plate splines, fields, Drunken Frenchman

The Mary Jane project includes an analysis of the slopes and other attributes of ski trails (Mary Jane/Winter Park Resort, CO) based on the USGS

high resolution digital elevation model. This study provides an illustration of the statistical tools in the R package, fields, for fitting smooth curves surfaces and displaying image data. Show cased here are methods of conditional simulation that provide useful measures of uncertainty for an estimated function.

Reconstruction of NH Average Temperature by Hierarchical Bayesian Models

❖ Bo Li, National Center for Atmospheric Research, 1300 30th Street Apt B3 24, Boulder, CO 80303, *boli@ucar.edu*; Douglas Nychka, National Center for Atmospheric Research; Caspar Ammann, National Center for Atmospheric Research

Key Words: BHM, bore-hole, NH temperature, tree ring

The past temperature is reconstructed based on proxies which are known to preserve the climate evolution. There are many kinds of proxies in term of temporal resolution (i.e., tree rings represent climate information at the interannual to decadal time scale, while bore-hole data record climate signals at multidecadal or longer time scales). We propose to integrate proxies of different time resolution by Hierarchical Bayesian Models (BHM) in reconstructing the past temperature. BHMs are built by three hierarchies which isolate the statistical errors of the proxy data from the underlying temperature process. We produce ensemble reconstructions of Northern Hemisphere (NH) average temperature for the past millennium based on tree ring and borehole based proxies.

Choice of Basis Functions in GAM Models of Air Pollution and Respiratory Health

❖ Chava Zibman, The University of Chicago, Department of Statistics, 5734 S University Ave, Chicago, IL 60637, *chava@uchicago.edu*; Vanja Dukic, The University of Chicago

Analyses of the relationship between pollution and morbidity over time generally rely on Poisson regression models with smooth functions of time to account for unobserved time-varying confounders. We propose an approach here, based on making data-based choice of bases for the smooth function, as an alternative to the usual spline bases. In particular, we illustrate the strategy using a set of trigonometric bases chosen according to the spectral density of the data. In a simulation study, we find that the performance of our method, as measured by the Mean Squared Error of the parameter of interest, is less sensitive than regression splines to the number of basis functions chosen and the extent to which the confounded and confounding variable vary quickly or slowly with time. We apply our model in an analysis of daily time-series asthma-prescription data over four summers in Chicago.

Statistical Approaches to El Niño Forecasting

❖ Souparno Ghosh, Texas A&M University, 415 college main street, Apt 6, College Station, TX 77840, *sghosh@stat.tamu.edu*; Amanda S. Hering, Texas A&M University; Salil Mahajan, Texas A&M University; Marc Genton, University of Geneva; Mikyoung Jun, Texas A&M University; Bani Mallick, Texas A&M University; Ramalingam Saravanan, Texas A&M University

Key Words: ENSO, MEOFs, Vector Autoregression, Model Selection, Bayes Factor, Credible interval

Predicting the evolution of climate on timescales of a season to a few years is challenging. The numerical approach tackles this problem using a supercomputer to solve the physical equations that govern the time evolution of the climate system. The statistical approach involves fitting empirical models to historical data. We construct a family of statistical models focusing on short-term forecasts of El NiÒo. Three types of models are fitted using Bayesian and a Classical methods: a simple VAR model; a seasonal monthly model; and a seasonal monthly model fit with a window of three months of data. We investigate the performance of these three models trained on monthly data during the period 1960–96 and tested during the period 1997-04.The statistical approach remains competitive with the numerical approach for El Nino forecasting, but is limited by the shortage of historical data.

An International Comparison of Air Quality Standards

❖ Peter Guttorp, University of Washington, Box 354322, Seattle, WA 98195-4322, *peter@stat.washington.edu*; Laura Knudsen, University of Washington

Key Words: Air quality, Environmental standards

Air quality standards vary among different countries of the world. While some nations, such as the United States, support their standards with enforcement mechanisms, other countries do not integrate implementation regulations into their policies. Furthermore, while the World Health Organization suggests health-based air quality standards, no international norms exist that regulate air quality. This discrepancy among different nations and their air quality standards forms the basis of our research. We have compiled a list of the air quality standards for different pollutants of different nations and regional bodies such as the European Union. The stringent World Health Organization standards are also included in this comparison. Additionally, we have compiled a comprehensive analysis of the enforcement policies for the different countries.

Space-Time Bayesian Environmental Modeling of the Chronic Wasting Disease in Deer

❖ Hae-Ryoung Song, University of South Carolina, 206 S Gregg Street Apt 6A, Columbia, SC 29205, *hrsong@gwm.sc.edu*; Andrew B. Lawson, University of South Carolina; Dennis Heisey, U.S. Geological Survey; Damien Joly, Alberta Sustainable Resource Development

We propose a Bayesian hierarchical model to investigate the pattern of spatial and temporal variation in disease prevalence of chronic wasting disease (CWD) in white-tailed deer in Wisconsin. The aims of this study are to describe the geographical distribution of CWD and to assess the effect of demographic factors such as age and sex on prevalence of CWD. Since we do not observe the exact infection time, we consider our data censored and develop survival models based on the imputed infection time. Our model explains the variation of CWD using individual covariates, while accounting for purely spatial, temporal and space-time interacted random effects. It also allows the investigation of space-time spread of the disease by incorporating spatial, temporal, and space-time interacted infection rates. We apply our model to white-tailed deer data collected in Wisconsin from 2001–2006.

Source Apportionment and Health Effects of Air Pollution

❖ Thomas Lumley, University of Washington, Department of Biostatistics, 1075 Pacific St N., Seattle, WA 98195, *tlumley@u.washington.edu*

Key Words: time series, pollution, factor analysis, receptor model

Particulate air pollution comes from a variety of sources that have different physical, chemical, and political characteristics. Attributing health effects of pollution to specific sources would be very valuable. Unfortunately source contributions are not directly measurable and must be imputed. I will discuss the source apportionment (receptor) models used for imputation and the problems of inference about source- specific health effects.

The Exploration of Space-Time Data

❖ Li Chen, The University of Chicago, The Center for Integrating Statistical and Environmental Science, 5734 S Ellis Ave Rm 459, Chicago, IL 60637, *lichen@uchicago.edu*; Michael Stein, The University of Chicago

Key Words: data exploration, space-time dependency

The typical data sets from geophysical and environmental sciences are rich in time and space. In order to model this kind of data statistically, the exploration of the space-time data is needed first. In this poster, we present some data exploration methods to study not only the trend surface but also the structure of space-time dependency. An application to air pollution data is described in depth.

186 Quantifying and Hedging Operational Risk ●

Section on Risk Analysis

Monday, July 30, 2:00 pm–3:50 pm

Quantifying Project Management Delay Costs Through Simulation Modeling

❖ Andrew Kumiega, Spectrum Global, 233 Maple, Downers Grove, IL 60515, *kumiega@comcast.net*; Deborah Cernauskas, IBM

Key Words: Simulation, Project Management

Most contractors have experienced time and cost overruns on a project. Unplanned project delays resulting in cost overruns occur for various reasons including material shortages and uncontrollable weather. Project managers are faced with the task of scheduling the activities needed to complete the project on time and on budget. At each step of the project, the possibility of a delay and cost overrun exists. The probability of a delay is generally taken into account by padding the duration of activities and cost estimates. The focus of this article is to allow the possibility of delays (non-excusable) in the development of the project critical path through simulation modeling. We also quantify the cost of crashing the stochastic CPM as well as determine the distribution of the resulting costs.

Environmental Risk Mitigation and Corporate Competitiveness

❖ Nasrin Khalili, IIT Stuart School of Business, 565 W Adams Street, Fourth Floor Room 420, Chicago, IL 60661, *khalili@iit.edu*

Key Words: Environmental Management, Risk Analysis, Risk Mitigation, Operation Strategies, Corporate Competitiveness, Environmental Risks

Corporations recognize the inevitability for managing environmental risks associated with the industrial economy. It is also evident that these risks must be mitigated in order to succeed in the competitive market. Mitigation strategies have been focusing on identifying environmental concerns within the realm of corporate strategy and development of codified management, audit, and reporting schemes for a wide range of the operations. This paper analyzes environmental risks and mitigation strategies at the leading organizations and describes how environmental decision making could feed into the entire value- adding process of corporations.

A Methodology for Trading and Investment System Development

❖ Benjamin Van Vliet, Illinois Institute of Technology, 1315 W Waveland, Chicago, IL 60613, *bvanvliet@stuart.iit.edu*; Andrew Kumiega, Spectrum Global

Key Words: Operational Risk, Financial markets, Development Methodology

Basel II defines operational risk as the probability of loss due to a failure of internal processes or external events. Best practices in software development lower operational risk. This paper presents best practices in trading and investment system development. The methodology combines a traditional development methodologies in a single paradigm with four stages. The activities of the four stages are organized into a plan-benchmark-do-check framework revolving around quantitative methods, data and testing, technology, and ongoing management respectively. The methodology incorporates gates from new product development, which are meetings where management makes go/kill decisions, weeding out weak projects and reallocating scarce resources towards more promising projects. We believe this methodology can be applied to software development-related problems in other industries.

Vendor Bid Evaluation Using a Risk-Adjusted Price Methodology

❖ Deborah Cernauskas, IBM, 4541 Saratoga Ave, Downers Grove, IL 60515, *dcernauskas@wans.net*

Key Words: risk adjusted price, bid evaluation

The bankruptcy of a service provider is an operational risk faced by any firm outsourcing business functions. The firm with the lowest bid is not necessarily the firm with the least risk. This article focuses on developing a risk adjusted price comparison methodology for multiyear contracts. A risk adjustment methodology for counterparty risk is developed using market expectations of service provider default risk. The risk adjusted price enables a comparison of the expected contract price versus the nominal price.

Project Portfolio Optimization for Operational Risk Management

❖ Aberrahim Labbi, IBM/Zurich Labs, Saeumerstr 4, 8803 Rueschlikon, Zurich, 8803 Switzerland, *abl@zurich.ibm.com*

Most businesses have to adapt strategies and operations to deal with various agents of change, such as customer requirements, technology evolution, etc. Business decisions under such constraints require taking some risks. Effective decision making requires extensive use of adaptive models which exploit various data sources for risk identification, assessment, & mgmt. In this paper we discuss some key statistical models for Operational risk mgmt and focus especially on IT projects risk analysis. We present an adaptive system for project portfolio risk assessment using a combination of statistical & machine learning techniques for estimation, simulation, & portfolio optimization. The system consists of a series of processes which can be summarized in three steps: risk factor identification and scoring; probabilistic estimation of project health; & project portfolio optimization.

187 Model Selection and Applications ● ✪

Biometrics Section, ENAR

Monday, July 30, 2:00 pm–3:50 pm

Adaptive Regressograms

❖ Anthony Gamst, University of California, San Diego, Biostatistics, Mail Code 0717, La Jolla, CA 92093-0717, *acgamst@math.ucsd.edu*

Key Words: minimax, rate adaptive, regression, penalized least squares, quasi-likelihood, model selection

We study the problem of selecting the number and location of break-points for piece-wise constant regression function estimates. When the underlying regression function has smoothness parameter s, it is known that the optimal number of bins increases at the rate 1/(2s+1). Simple model selection penalties lead to regressograms with nearly optimal risk. These estimates also adapt to the unknown smoothness s, leading to optimal rates of convergence regardless of the underlying smoothness. The arguments involved can be extended to other spline bases and to quasi-likelihood-type regression problems.

On L1-Norm Multiclass Support Vector Machines: Classification of High-Dimension, Low Sample Size Data

❖ Lifeng Wang, University of Pennsylvania, 624 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, *iamwlf@gmail.com*; Xiaotong Shen, The University of Minnesota

Key Words: High-dimension and low sample size, Margin classification, Regularization, Sparsity, Variable selection

Binary support vector machines (SVM) have proven to deliver high performance. In multiclass classification, issues remain with respect to variable selection. A challenging issue is classification and variable selection in the presence of variables in the magnitude of thousands, greatly exceeding the size of training sample. This often occurs in genomics classification. We propose a novel multiclass SVM, which, together with a developed regularization solution path, perform classification and variable selection simultaneously through an L1-norm penalized sparse representation. A statistical learning theory is developed to quantify the generalization error in an attempt to gain insight into the basic structure of sparse learning, permitting the number of variables to greatly exceed the sample size. The numerical results suggest that the proposed methodology is highly competitive.

Fence Methods for Quantitative Trait Loci Mapping in Humans

❖ Thuan Nguyen, University of California, Davis, 2900 Solano Park Circle Apt 2623, Davis, CA 95616, *tnguyen@wald.ucdavis.edu*; Jiming Jiang, University of California, Davis; Jie Peng, University of California, Davis

Key Words: Fence method, Identity by descent (IBD), Model selection, Pedigree data, QTL mapping, Variance component

The goal of quantitative trait loci (QTL) mapping is to identify chromosomal regions that affect particular quantitative traits. In humans, pedigree data is collected to detect the cosegregation of phenotypes and the shared genetic material between relatives. In QTL mapping, variance components arising from the trait genes, polygenic and environmental effects are often used to model the covariance structure of the phenotypes given the IBD sharing matrix. The genetic control of a quantitative trait is widely thought to involve a number of loci. However, standard genome scanning methods usually ignore these possibilities, presumably because they involve more complex models partially due to the large number of putative trait loci. Thus, a model selection approach seems useful in such a context. We explore a recently developed model selection strategy, called fence method, in QTL mapping.

Model Selection Under the Proportional Hazards Mixed Effects Model (PHMM)

❖ Michael Donohue, University of California, San Diego, 9500 Gilman Drive MC 0717, San Diego, CA 92093-0717, *mdonohue@ucsd.edu*; Ronghui Xu, University of California, San Diego; Anthony Gamst, University of California, San Diego; Florin Vaida, University of California, San Diego; David P. Harrington, Dana-Farber Cancer Institute

Key Words: model selection, proportional hazards, mixed effects, Akaike information criterion

In this talk we discuss model selection under the proportional hazards mixed effects model (PHMM). Recently established asymptotic properties of the nonparametric maximum likelihood estimator allow us to use the profile likelihood for selection of both nested and non-nested PHMMs. We define a profile Akaike information for general models with nuisance parameters. Asymptotic quadratic expansion of the log profile likelihood allows unbiased estimation of the Akaike information by a profile Akaike information criterion (pAIC). The pAIC focuses on the population parameters, such as the fixed effects and variance components. We will also discuss conditional AIC (cAIC), which, in turn, incorporates the estimated random effects. Computation of both pAIC and cAIC under PHMM will be addressed, and examples will be given to show their applications.

Model Selection for Partial Spline Models

❖ Hao Zhang, North Carolina State University, 2501 Founders Drive, 216 Patterson Hall, Raleigh, NC 27695-8203, *hzhang2@stat.ncsu.edu*

Key Words: LASSO, oracle, model selection, shrinkage, partial spline

Partial spline models provide good compromises between linear and nonparametric models. Due to model complexity, model selection is challenging for partial spline models and consists of two problems: smoothing parameter selection for the nonparametric part and variable selection for parametric covariates. We propose a unified approach for automatic model selection in partial spline models via a double-penalized likelihood procedure, where an adaptive soft-thresholding penalty is used for selecting important parametric effects. We show that, if the smoothing and regularization parameters are properly chosen, the procedure is as efficient as the oracle estimator (Fan, 2001). Numerical results are presented as well.

188 Statistical Challenges in Evolutionary Biology ● ✪

IMS

Monday, July 30, 2:00 pm–3:50 pm

Analysis of Count Data with Phylogenetic Random Effects

❖ Cecile Ane, University of Wisconsin-Madison, Department of Statistics, 1300 University Ave, Madison, WI 53706, *ane@stat.wisc.edu*

Key Words: evolution, phylogenetic effect, mixed model, simulated maximum likelihood

I will describe a statistical method for analyzing count data observed over varying times and multiple individuals from a number of related species. This method accounts for random phylogenetic effects that cause correlation in observations. This generalized linear mixed model regression approach is based on simulated maximum likelihood. It will be used to evaluate the effects of altitude, nectar reward and flower display on the frequency of hummingbird visitation.

Coalescent-Based Inference of Population Dynamics with Gaussian Markov Random Field Temporal Smoothing

❖ Vladimir N. Minin, University of California, Los Angeles, Biomathematics, 695 Charles E. Young Dr., Los Angeles, CA 90095-1766, *vminin@ucla.edu*; Marc A. Suchard, University of California, Los Angeles

Key Words: coalescent, effective population size, smoothing, Gaussian Markov Random Field

Kingman's coalescent process opens the door for estimation of population genetics model parameters from molecular sequences. One paramount parameter of interest is the effective population size. Temporal variation of this quantity characterizes the demographic history of a population. We propose to exploit Gaussian Markov random fields (GMRFs) to achieve temporal smoothing of the effective population size in a Bayesian framework. In a simulation study, we demonstrate that the proposed temporal smoothing method successfully recovers "true" population size trajectories in all simulation scenarios. We apply our GMRF smoothing to sequences of hepatitis C virus contemporaneously sampled in Egypt and human influenza A hemagglutinin sequences serially sampled throughout three flu seasons.

A Step Toward Bar Coding Life: A New Method To Assign Genes to Pre-existing Species Groups

❖Zaid Abdo, University of Idaho, Dept of Mathematics Brink 300, Moscow, ID 83844-1103, *zabdo@uidaho.edu*; Geoffre B. Golding, McMaster University

Key Words: barcoding, assignment, decision theory, coalescent

A major part of the bar coding of life problem is to be able to assign newly sequenced or sampled individuals to existing groups that are pre-identified externally (by a taxonomist, for example). This problem involves evaluating the statistical evidence towards associating a new individual with a group or another. The main concern of our current research is to perform this task in a fast and accurate manner. To accomplish this we developed a model-based, decision theoretic framework based on the coalescent theory. Under this framework, we utilize both distance and the posterior probability of a group given the data and the newly sampled individual to assign this new individual. We believe this approach maximizes the use of the available information in the data. Our preliminary results indicate this approach is superior to using a simple measure of distance for assignment.

Measures of Volatility and Valency for Serially Sampled Nucleotide Sequences

❖John O'Brien, University of California, Los Angeles, Box 951766 - Biomathematics, Medical School, Los Angeles, CA 90095-1766, *jdobrien@ucla.edu*

Key Words: sequence volatility, serially sampled, Markov process, phylogenetics

(Plotkin and Dushoff, 2004) showed that it is possible to define a sequence volatility measure, giving the number of amino acid changes accessible within one nucleotide substitution of a given codon. We define a related quantity, the valency of codon, which is the volatility modulo the number of different residues available. Starting with the Goldman/Yang matrix for codon changes, we partition the process via a Markov counting process (Minin, 2007) and are able to derive pairwise distances for both valency and volatility. In the case of influenza A hemagglutinin, we show that after zoonotic transfer the valency of the sequence increases substantially over time, with the most dramatic alterations at known active sites in the protein. We conclude by showing that these changes are independent of shifts in the underlying nucleotide distribution.

Phylogenies Unplugged: Consensus Trees with Wandering Taxa

❖Benjamin Redelings, North Carolina State University, 4305 Avent Ferry Rd Apt 2, Raleigh, NC 27606, *benjamin_redelings@ncsu.edu*

Key Words: Bayesian, evolution, phylogeny, tree, consensus, wandering

Evolutionary tree topologies are discrete and unordered and so their posterior distributions cannot be summarized by a mean and variance. One common summary, the majority tree, represents the "center" of a topology distribution by combining supported taxa bi-partitions into a single tree. However, if a clade wanders widely over the topology of the remaining branches, the majority tree may collapse to a star tree that contains no information. I present an improved summary that can reveal hidden structure in the posterior distribution by representing wandering clades instead of removing leaf taxa. The summary is a generalization of multifurcating (MF) trees called a multi-connected (MC) tree. An MC tree can be represented as a collection of bi-partitions of leaf taxon subsets and has a useful visual representation as a graph containing all trees which contain these bi-partitions.

189

Statistics and Machine Learning in High Dimension ● ✪

Section on Bayesian Statistical Science

Monday, July 30, 2:00 pm–3:50 pm

Estimating Rates of Rare Events in Massive Web Applications

❖Deepak Agarwal, Yahoo! Research, 2825 Mission College Blvd, Santa Clara, CA 95054, *dagarwal@yahoo-inc.com*

We consider the problem of estimating occurrence rates of rare events for massive web scale applications. In particular, we focus on the problem of estimating click-through rates for (query, ad) pairs in the context of internet advertising. We consider a scenario where both queries and ads are classified into taxonomies that capture broad contextual information at different levels of granularity. We provide a two stage model that exploits the hierarchical structure induced by the taxonomies and provide accurate estimates at multiple resolutions. Our first stage model is based on a novel imputation strategy that estimates the denominator of the rates. The model also adjusts for the sampling bias introduced due to crawling constraints in our application. Conditional on the first stage estimates, our second stage model uses a tree-structured Markov model to provide smooth rate estimates.

Robust Test for Detecting a Signal in a High-Dimensional Sparse Normal Vector

❖Junyong Park, University of Maryland, Baltimore County, Department of Mathematics and Statistics, 1000 hilltop circle, Baltimore, MD 21250, *junpark@math.umbc.edu*; Eitan Greenshtein, Statistical and Applied Mathematical Sciences Institute

Key Words: high dimension, robust test

Let $Z_i, i=1,...,n$ independent random variables with mean μ_i and variance 1. We consider the problem of testing $H_0: \mu_i=0, i=1,...,n$. The setup is when n is large, and the mean vector is `sparse.' We suggest a test which is not sensitive to the exact tail behavior implied under normality assumptions. In particular, if the 'moderate deviation' tail of the distribution of Z_i, may be represented as the product of a tail of a standard normal and a 'slowly changing' function, our suggested test is robust. Such a tail behavior, and a need for such a robust test, is expected when the Z_i are of the form $Z_i=\sum_{j=1}^m Y_{ij}/ \sqrt{m}$, for large m, $m< < n$, and independent Y_{ij}.

Learning Gradients and Feature Selection on Manifolds

❖Sayan Mukherjee, Duke University, 101 Science Drive, Durham, NC 27708, *sayan@stat.duke.edu*

Key Words: Manifold learning, Dimensionality reduction and regression, Kernel models

An underlying premise in the analysis and modeling of high-dimensional physical and biological systems is that data generated by measuring thousands of variables lie on or near a low-dimensional manifold. This premise has led to various estimation and learning problems grouped under the heading of ``manifold learning." It is natural to formulate the problem of feature selection—finding salient variables (or linear combinations of salient variables) and estimating how they covary—in the manifold setting. For regression and classification the idea of selecting features via estimates of the gradient of the regression and classification function has been developed. In this paper we extend this approach from the Euclidean setting to the manifold setting.

Feature Extraction for Classification of Functional Data

❖ Bin Li, Louisiana State University, 161 Ag Admin Bldg, Louisiana State University, Baton Rouge, LA 70803, *bli@lsu.edu*; Qingzhao Yu, Louisiana State University Health Sciences Center

Key Words: classification, feature extraction, functional data

Irregular functional data, characterized by spatial heterogeneity and local features like peaks, are commonly encountered in biomedical research, such as spectrometric data. However, most of the current methods in functional data analysis are based on smoothing methods using global penalties and bandwidth. In this presentation, we proposed a novel feature extraction method which is particularly useful in detecting local features for classification of functional data. More specifically, it wisely selects a few "markers" over the spectrum and utilizes the discriminant coordinates on the subspace around these markers as the extracted features. We apply the method to two public domain data sets together with simulated data and discuss the understanding developed from these applications.

Dimension Augmenting Vector Machine: A New General Classifier System for Large p Small n Problem

❖ Samiran Ghosh, Indiana University Purdue University Indianapolis, 402 N Blackford Street LD270, Indianapolis, IN 46202, *samiran@math.iupui.edu*; Yazhen Wang, University of Connecticut; Dipak Dey, University of Connecticut

Key Words: Classification, Import Vector Machine, Radial Basis Function, Regularization, Reproducing Kernel Hilbert Space, Support Vector Machine

Support vector machine and other reproducing kernel Hilbert space based classifier systems are drawing much attention recently due to its robustness and generalization capability. All of these approaches construct classifier based on training sample in a high dimensional space by using all available dimensions. SVM achieves huge data compression by selecting only few observations lying in the boundary of the classifier function. However when the number of observations are small but the number of dimensions are very large then it is not necessary that all available dimensions are carrying equal information in the classification context. Selection of only useful fraction of available dimensions will result in huge data compression. In this paper, we have come up with an approach parallel to IVM introduced by Zhu et al., by means of which such an optimal set of dimension could be selected.

190 Section on Bayesian Statistics Student Paper Competition: Network Models, Reliability, and Prediction ● ✪

Section on Bayesian Statistical Science

Monday, July 30, 2:00 pm–3:50 pm

A Bayesian Mixed Effects Model for Longitudinal Social Network Data-Student Paper Competitions

❖ Anton Westveld, Washington University in St. Louis, Campus Box 1120, One Brookings Drive, St Louis, MO 63130, *westveld@stat.washington.edu*; Peter D. Hoff, University of Washington

Key Words: dyadic data, relational data, longitudinal, social networks, Bayesian estimation, regression modeling

This paper is concerned with a Bayesian approach to the estimation of models for data that arise from measurements made on pairs of actors, where every ordered pair of a group of actors is potentially measured at regular temporal intervals, resulting in social network data for each point in time (longitudinal social network data). Typically social network data are used to study a key social phenomenon, such as trade between nations, in relation to a set of predictor variables while accounting for and learning about the interconnectivity of the actors. The network and temporal dependencies are both based on Markov structures and are modeled through a random effects approach resulting in a stochastic process defined by a set of stationary covariance matrices. We apply the methodology to two real-world datasets: international trade and militarized interstate disputes.

Bayesian Network-Based Process Monitoring and Diagnosis in Complex Manufacturing Systems

❖ Jing Li, The University of Michigan, Ann Arbor, MI 48105, *jinglz@umich.edu*; Jianjun Shi, The University of Michigan; Jionghua Jin, The University of Michigan

Key Words: Bayesian network, causal model, Statistical Quality Control, Statistical Process Control, diagnosis

Process monitoring and diagnosis is a challenging problem in manufacturing systems with many variables and complex relationships. Because diagnoses require tracing backward from a fault to its root causes, the causal relationships among the variables must be identified. Bayesian networks provide a tool to model the causal relationships, enabling effective diagnoses. This paper proposes a "Bayesian Network based T2 decomposition" method, which uses the causal relationships, represented by a Bayesian Network, to guide the decomposition of the out-of-control signal in a T2 control chart for diagnostic purposes. Theoretical analysis and simulation studies demonstrate that the proposed method has a substantially enhanced diagnosability and reduced computational complexity, compared with the traditional approach that decomposes the out-of-control signal without prior causal identification.

The Dependent Poisson Race Model and Modeling Dependence in Conjoint Choice Experiments (Student Paper Competition)

❖ Shiling Ruan, The Ohio State University, 2001 Whitman Way, Apt 30, San Bruno, CA 94066, *ruan.8@osu.edu*; Steven N. MacEachern, The Ohio State University; Thomas Otter, The Ohio State University; Angela Dean, The Ohio State University

Key Words: Poisson race model, Bayesian model, multinomial logit model, dependent processes, dominance, Poisson process

Conjoint choice experiments are used widely in marketing to study consumer preferences amongst alternative products. We develop a class of choice models, belonging to the class of Poisson race models, that describe a ``random utility'' which lends itself to a process-based description of choice. The models incorporate a dependence structure which captures the relationship between the attributes of the choice alternatives and which appropriately moderates the randomness inherent in the race. The new models are applied to conjoint choice data and are shown to have markedly superior performance to independent Poisson race models and to the multinomial logit model.

Probabilistic Weather Forecasting for Winter Road Maintenance

❖Veronica J. Berrocal, University of Washington, Box 354322, Department of Statistics, Seattle, WA 98195, *veronica@stat.washington.edu*; Adrian E. Raftery, University of Washington; Tilmann Gneiting, University of Washington

Key Words: spatial hierarchical model, probabilistic weather forecasting, predictive probability distributions

Road maintenance is a critical issue during winter time. In order for anti-icing operations to be efficient, predictions of road ice need to be accurate and reliable. Probabilistic forecasts take forecast uncertainty into account and allow decision makers to make informed, flexible judgments. In this paper, we propose a spatial hierarchical model that post-processes numerical weather forecasts and yields joint predictive probability distributions of precipitation occurrence and temperature. These in turn lead forecasts of the probability of ice on the road. In predictions of ice formation along Washington State Interstate 90 Mountains to Sound Greenway for the 2003–2004 and 2004–2005 winter seasons, the statistically post-processed probabilistic forecasts were sharp and calibrated, and had higher economic value than the deterministic forecasts from a numerical weather prediction model.

191 Synthetic Data and Other Methods for Disclosure Limitation and Confidentiality Preservation

Section on Survey Research Methods, Social Statistics Section

Monday, July 30, 2:00 pm–3:50 pm

Microdata Simulation for Confidentiality Protection Using Regression Quantiles and Hot Deck

❖Jennifer Huckett, Iowa State University, 114 Snedecor, Ames, IA 50011, *jhuckett@iastate.edu*; Michael D. Larsen, Iowa State University

Key Words: disclosure control, disclosure avoidance, synthetic data, multiple imputation, tax return data

Government agencies must simultaneously maintain confidentiality of individual records and disseminate useful microdata. Iowa's Legislative Services Agency (LSA) needs predicted state tax revenue based on proposed policy changes calculated from individual income tax returns. Iowa's Department of Revenue (IDR) cannot provide individual records to LSA by law. Currently, LSA submits requests to IDR that IDR computes and reports to LSA. This is inefficient for both agencies. We study options for IDR creating a full synthetic tax return file for release to LSA. Specifically, we study combining quantile regression, hot deck imputation, and additional confidentiality-preserving methods to produce releasable, usable data. Several versions of microdata can be multiply imputed to assess uncertainty. Measures of disclosure risk to evaluate confidentiality protection are considered.

Releasing Multiply Imputed, Synthetic Data Generated in Two Stages To Protect Confidentiality

❖Joerg Drechsler, Institute for Employment Research, Weddigenstrasse 20-22, Nuremberg, 90478 Germany, *joerg.drechsler@iab.de*; Jerome P. Reiter, Duke University

Key Words: Confidentiality, Disclosure, Multiple Imputation, Synthetic Datasets

In 1993 Rubin suggested to generate synthetic datasets from existing survey data if a release of the unaltered original data is impossible for confidentiality reasons. However, generating datasets that preserve inferences for a large number of possible estimates is a difficult and labor intensive task. Furthermore, considering the partially synthetic data framework, the release of too many different imputations for sensitive variables that vary only slightly might enable the intruder to guess the true values for these variables very closely. To reduce this risk and the burden of generating synthetic datasets, a two stage imputation process can be used. In this paper, we introduce this new approach, first developed for a German establishment panel. We derive new combining rules for the synthetic datasets necessary in this context and evaluate the new estimates with a simulation study.

Multiple Imputation Alternatives to Top-Coding for Statistical Disclosure Control

❖Di An, University of Michigan, Department of Biostatistics, 1420 Washington Heights, Ann Arbor, MI 48109, *dianch@umich.edu*; Roderick J. Little, University of Michigan

Key Words: confidentiality, disclosure protection, multiple imputation

Top-coding of extreme values of variables like income is a common method of statistical disclosure control, but it creates problems for the data analyst. This article proposes two alternative methods to top-coding for SDC based on multiple imputation (MI). We show in simulation studies that the MI methods provide better inferences of the publicly-released data than top-coding, using straightforward MI methods of analysis, while maintaining good SDC properties. We illustrate the methods on data from the 1995 Chinese household income project.

Secure Logistic Regression

❖Yuval Nardi, Carnegie Mellon University, 5921 Douglas St, Pittsburgh, PA 15217, *yuval@stat.cmu.edu*; Stephen Fienberg, Carnegie Mellon University; Aleksandra Slavkovic, The Pennsylvania State University

We address the problem of performing a logistic regression in a secure way without directly sharing data among parties. We suppose that data are collected separately and intimately by several separate parties (agencies). These parties wish to analyze the pooled (combined) data without actually combining the parts that they possess (i.e., they want to fit a model and make inferences using the pooled data in a way that no one party's data are disclosed to another party). In this paper we build on earlier results by Fienberg, Fulp, Slavkovic and Wrobel (2006) on the horizontally partitioned case and describe methods for both the vertically partitioned case and the general case, the vertically partitioned, partially overlapping data.

Application of the Truncated Triangular and the Trapezoidal Distributions for Developing Multiplicative Noise

❖Jay J. Kim, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *pzk3@cdc.gov*

Key Words: truncated triangular distribution, masking, confidentiality, trapezoidal distribution

It has been suggested that the truncated triangular distribution be used for masking microdata. The random variable which follows the truncated triangular distribution serves as a multiplicative noise factor. The natural candidate distribution is the one which is centered at 1, but truncated symmetrically around 1, because it is not desirable to use some values which are too close to 1 due to confidentiality concerns. Instead of truncating the middle section of the distribution, we can assign low probabilities for that part of the distribution. This procedure will render some variations of the trapezoidal distribution. A third approach would be to use the upside-down triangular distribution for the middle section of the distribution. In this paper, we will derive the probability density function for each of the three distributions and investigate their properties.

192 Modeling Longitudinal Data in Epidemiology ●

Section on Statistics in Epidemiology, Biometrics Section, ENAR, Section on Health Policy Statistics, Section on Teaching Statistics in the Health Sciences

Monday, July 30, 2:00 pm–3:50 pm

Computationally Efficient Estimation of Multilevel High-Dimensional Latent Variable Models

❖ Bengt Muthen, University of California, Los Angeles, 3463 Stoner Ave, Los Angeles, CA 90066, *bmuthen@ucla.edu*; Tihomir Asparouhov, Muthen & Muthen

Key Words: Categorical variables, Weighted least squares, Twin analysis

Multilevel analysis often leads to modeling with multiple latent variables on several levels. While this is less of a problem with Gaussian observed variables, maximum-likelihood (ML) estimation with categorical outcomes presents computational problems due to multi-dimensional numerical integration. We describe a new method that compared to ML is both computationally efficient and has similar MSE. The method is an extension of the Muthen (1984) weighted least squares (WLS) estimation method for multilevel multivariate latent variable models for any combination of categorical, censored, and normal observed variables. Using a new version of the Mplus program, we compare MSE and the computational time for the ML and WLS estimators in a simulation study and present a longitudinal example where heritability is estimated using MZ and DZ twins.

Interval-Censored Regression for Modeling the Effect of Labor Progression on Uterine Rupture

❖ Sarah Ratcliffe, University of Pennsylvania, 619 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19038, *sratclif@cceb.upenn.edu*; George A. Macones, Washington University in St. Louis

Key Words: interval censoring, longitudinal

During the first stage of labor, a woman's progression is routinely monitored by cervical dilation. The dilation over time is known as the labor curve, and the timing of each centimeter increase in dilation can yield potentially useful information about labor complications. However, women may be at any dilation when they arrive at the hospital and dilations are only monitored intermittently during labor. Thus, the exact timing of each transition from dilation to the next is unknown. Rather, we have a time interval over which the transition occurred. We show how interval censored regression can be used to examine the effect of labor progression on the risk of uterine rupture in women attempting a vaginal birth after a cesarean section (VBAC).

Longitudinal Latent Variable Models in Environmental Epidemiology

❖ Brisa Sanchez, University of Michigan, Biostatistics Department, 1420 Washington Heights 4th Fl, Ann Arbor, MI 48109, *brisa@umich.edu*

Key Words: latent variables, design, lead

We use structural equation modeling with latent variables to study the relationships between bone- and blood-based biomarkers of fetal lead exposure using data from a cohort study conducted in Mexico City. Prospective mothers were ideally recruited during or before the first trimester of pregnancy, and followed until birth. Various blood-based lead exposure biomarkers were measured repeatedly at different points during gestation, and umbilical cord blood lead measurements were obtained at birth. Measurements of cortical and trabecular bone lead concentrations were obtained before pregnancy and one month after delivery. The proposed model succinctly describes the associations of interest and utilizes all data available, as opposed to conducting multiple analyses with traditional regression methods.

Bayesian Multivariate Growth Curve Latent Class Models for Mixed Outcomes

❖ Benjamin E. Leiby, Thomas Jefferson University, 1015 Chestnut St, Suite M100, Philadelphia, PA 19107, *bleiby@mail.jci.tju.edu*; Mary D. Sammel, University of Pennsylvania; Thomas R. Tenhave, University of Pennsylvania; Kevin G. Lynch, University of Pennsylvania

Key Words: Latent Variable, Latent Class, Multivariate, Longitudinal, Growth Mixture

In studies of complex diseases, multiple outcomes are often used to adequately capture information about disease severity. The disease population may also have distinct subgroups and identification of these subgroups is of interest as it may assist clinicians in providing appropriate treatment or in developing accurate prognoses. We propose Bayesian models that group subjects based on multiple continuous, binary, ordinal or count outcomes measured repeatedly over time. These groups or latent classes are defined by distinctive longitudinal profiles of a latent variable which is used to summarize the multivariate outcomes at each point in time. The mean growth curve for the latent variable in each class defines the features of the class. We apply the model to data from a clinical trial evaluating the efficacy of Bacillus Calmette-Guerin in treating symptoms of Interstitial Cystitis.

Modeling Longitudinal Steroid Hormone Levels for Predicting Menopausal Symptoms

❖ Mary D. Sammel, University of Pennsylvania, 707 Harvey Road, Wallingford, PA 19086, *msammel@mail.med.upenn.edu*; Jeanne Manson, Children's Hospital of Philadelphia; Ellen W. Freeman, University of Pennsylvania

Key Words: change, Longitudinal Profiles, mixed models

In a population-based cohort designed to study hormonal effects of ovarian aging, we estimated hormone trajectory and variability of each participant over four years. We evaluated whether these components predicted menopausal hot flashes, and compared these model-based estimates of estradiol (E) and follicle-stimulating hormone (FSH) with standard statistical summaries. We hypothesized that change in hormone levels rather than absolute levels predicted menopausal symptoms. The quadratic estradiol trend (reflecting increasing then decreasing levels over four years) and the maximum observed FSH significantly predicted the prevalence and incidence of hot flashes. These results demonstrate that it is the trajectory of the hormone profile and not the absolute hormone level that is associated with menopausal hot flashes.

193 Smoothing Spline Methods and Their Applications

WNAR, Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Nonparametric Hazard Estimation for Gap Time in Recurrent Event Data

❖ Pang Du, Virginia Polytechnic Institute and State University, 410 Hutcheson Hall, Department of Statistics, Blacksburg, VA 24061-0439, *pangdu@vt.edu*

Key Words: Recurrent event, Gap time, Hazard, Penalized likelihood, Asymptotic convergence rate, Model Selection

Recurrent event data arise in many biomedical and engineering studies when failure events can occur repeatedly over time for each study subject. In this article, we are interested in nonparametric estimation of the hazard function for gap times. A penalized likelihood model is proposed to estimate the hazard as a function of both gap time and covariate. Method for smoothing parameter selection is developed and Bayesian confidence intervals for the hazard function are derived. Asymptotic convergence rates of the estimates are also established by assuming no gap times of a subject are the same. Empirical studies are performed to evaluate various aspects of the method. The exploratory role of the proposed technique is illustrated through an application to the well-known bladder tumor cancer data.

Spectral Density Estimation via State Space Modeling

❖ Li Qin, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave North LE400, Seattle, WA 98109, *lqin@scharp.org*; Dongfeng Li, Fred Hutchinson Cancer Research Center

Key Words: Locally-stationary process, Simulation smoother, Smoothing spline, Spectral analysis, State space model, Time series

We consider periodogram smoothing for the spectral density estimation of a stationary or non-stationary time series. Smoothing splines are used to model the periodograms which permits a connection with an equivalent state space formulation. An approximate Gaussian state space model and the associated algorithm are derived to carry out the estimation, inference and smoothing parameter selection in a unified framework. The proposed methods are demonstrated by examples with simulated and real time series.

Pointwise Consistency of Smoothing Spline Estimate

❖ Ming Dai, The University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28213, *mdai@email.uncc.edu*; Wensheng Guo, University of Pennsylvania

Key Words: nonparametric estimate, point-wise consistent, smoothing spline estimate

In this paper, we develop new methods for studying the asymptotic properties of smoothing spline estimate. Because of the difference between the smoothing spline estimate as a global estimate and the local estimates, we propose to approximate the estimate directly with certain basis, instead of approximating the weight function in smoothing spline estimate. With the approximation, we show the different asymptotic properties of smoothing spline estimate as a global estimate from the local estimates. With uncorrelated errors, we show that the smoothing spline estimate from the penalized least square is point-wise consistent, and achieves the optimal convergence rate. With short- and long-memory errors, we show that the smoothing spline estimate from the penalized weighted least square is also point-wise consistent, and achieves the optimal convergence rate.

M-Type Smoothing Spline ANOVA for Correlated Data

❖ Anna Liu, University of Massachusetts Amherst, Amherst, MA 01003, *anna@math.umass.edu*; John Staudenmayer, University of Massachusetts Amherst; Li Qin, Fred Hutchinson Cancer Research Center

Key Words: Robust, Smoothing spline ANOVA, Correlated data, Resistant smoothing parameter, Robust inference, Hypothesis test

This paper extends the M-type smoothing spline regression models by Huber (1979) for independent observations with a single smoothing parameter to smoothing spline ANOVA models for correlated data. Simultaneous resistant estimates of the parameterized covariance matrix and the smoothing parameters are developed. This framework allows robust inference on the regression function. Inference procedures are developed for general hypotheses and evaluated through simulations for testing a constant regression function and no difference of mean group curves. Simulations show that the robust tests have improved performances over their Gaussian counterparts. Applications to real data will be demonstrated.

Penalized Functional Principal Components Analysis Using a Kullback-Leibler Criterion

❖ Robert Krafty, University of Pennsylvania, 423 Guardian Drive, 501 Blockley Hall, Philadelphia, PA 19104, *krafty@mail.med.upenn.edu*; Wensheng Guo, University of Pennsylvania

Key Words: Functional Data Analysis, Functional Principal Components Analysis, Penalized Estimation, Smoothing Parameter Selection

A new penalized method for preforming functional principal components analysis is proposed. The penalty is based on a regularity of the underlying covariance function that is characterized by the balance between the decays of the eigenvalues and smoothness of the eigenfunctions. We propose to use the Kullback-Leibler distance to directly measure the goodness-of-fit in the space of functional covariances. A leave-one-trajectory-out cross validation procedure is introduced to estimate this Kullback-Leibler distance that allows for the selection of the smoothing parameter and the number of principal components. Asymptotic consistency rates are computed which show that procedures selecting smoothing parameters based on metrics over the space of subject trajectories behave sub-optimally and the method is used to analysis the time course gene expression of fibroblast.

194 What Have We Learned from Results of Primary Prevention Cardiovascular Medical Device Mortality Trials? ● ✪

Biopharmaceutical Section, ENAR

Monday, July 30, 2:00 pm–3:50 pm

What Have We Learned from Results of Primary Prevention Cardiovascular Medical Device Mortality Trials?

❖ Kerry Lee, Duke University, 2400 Pratt Street, Room 0311 Terrace Level, Durham, NC 27705, *kerry.lee@duke.edu*; ❖ Laura Thompson, Food and Drug Administration, 1350 Piccard Dr HFZ550, CDRH, Rockville, MD 20850, *laura.thompson@fda.hhs.gov*; ❖ Andrew Mugglin, The University of Minnesota, 2221 University Ave SE, Suite 200, Minneapolis, MN 55414, *amugglin@umn.edu*; ❖ David Breiter, Boston Scientific Corporation, 4100 Hamline Ave N, St Paul, MN 55112, *david.breiter@guidant.com*; ❖ W. Jackson Hall, University of Rochester, Department of Biostatistics, School of Medicine and Dentistry, Rochester, NY 14642-8630, *hall@bst.rochester.edu*

Key Words: ICDs, primary-prevention, sequential, medical-device, endpoints, trials

Sudden cardiac death claims about Ω million lives every year in the United States. The MADIT I, II, and SCD-HeFT trials demonstrated that implantable cardioverter defibrillators (ICDs) reduce total mortality in patients who have a high risk of experiencing SCD. These landmark trials raised many interesting statistical questions. Statistical methods for analyzing survival data and fully sequential (MADIT and MADIT-II) group sequential designs (SCD-HeFT) were employed. In addition to total mortality, trial endpoints included health outcomes measurements (e.g., cost-effectiveness). In this session, panelists will discuss their perspectives on statistical, clinical, public policy and regulatory issues associated with the results and wide-scale acceptability of these trials and uptake of ICDs. Panelists will also provoke discussion of future directions for endpoints.

195 Career Opportunities for Biostatisticians: Perspectives from Industry, Government and Academics ● ✪

Biopharmaceutical Section, ENAR, Section on Statistical Consulting, WNAR

Monday, July 30, 2:00 pm–3:50 pm

Career Opportunities for Biostatisticians: Perspectives from Industry, Government, and Academia

❖ Tammy Massie, Food and Drug Administration, CBER, 1401 Rockville Pike, HFM217, Rockville, MD 20852, *tammy.massie@fda.hhs.gov*; ❖ Stacey Cofield, The University of Alabama at Birmingham, 1530 3rd Ave S RPHB 327, Birmingham, AL 35294-0022, *scofield@uab.edu*; ❖ Carol Summitt, Wyeth, PO Box 26609, Richmond, VA 23261, *summitc@wyeth.com*; ❖ Laura Stets, Food and Drug Administration, 7500 Standish Place, CVM, MPN II, Derwood, MD 20855, *laura.stets@fda.hhs.gov*

Key Words: Career Opportunities, Academia, Industry, Government

This session will have speakers from industry, government and academia. Within this session several speakers will present a brief background of their career track after receiving an advanced degree in biostatistics. Each speaker will illustrate and present observations related to choosing and developing within their chosen career field. In particular, the speakers will highlight typical daily tasks and activities, primary roles and responsibilities, opportunities for research and professional development. Each speaker will also illustrate expected and unexpected characteristics and benefits related to their career choice. Furthermore, suggestions related to career development for students (or professionals in other fields) to prepare for employment opportunities and to grow professionally will be discussed.

196 Recent Advances in Microarray Data Analysis II ● ✪

Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Model-Based Approach for Cancer Outlier Differential Gene Expression Detection

❖Baolin Wu, The University of Minnesota, 420 Delaware St SE, A460 Mayo Building MMC 303, Minneapolis, MN 55455, *baolin@biostat.umn.edu*

Key Words: Differential expression detection, K-means clustering, Least absolute deviation, Microarray, Outlier, Partition around medoids

We study model based approaches to detect cancer genes that are over/down-expressed in some but not all samples in a disease group. This has proven quite useful in cancer studies since heterogeneous oncogene activation patterns have been observed in the majority of cancer types. In this paper we propose the model based outlier differential expression (MODE) detection methods, which are designed to automatically identify cancer outliers and quantify the differential expressions simultaneously. The proposed MODE can be derived from the likelihood ratio test under the Gaussian and exponential error models, and efficient numerical algorithms are developed for the maximum likelihood ratio computation. Using real and simulation studies, we compare the proposed MODE to the existing outlier differential expression detection methods in the literature and illustrate its competitive performance.

Application of Discrimination Method on Combining Gene Expression Data of Kidney Transplant Patients

❖ Xiangrong Kong, Virginia Commonwealth University, 3524 Grove Ave Apt 7, Richmond, VA 23221, *kongx@vcu.edu*; Kellie J. Archer, Virginia Commonwealth University

Key Words: chronic allograft nephropathy, microarray, discrimination, kidney tranplant, gene expression

Chronic allograft nephropathy (CAN) is a major cause of graft loss and morbidity among patients with kidney transplantation. It may be promising to characterize the gene expression pattern of CAN and use it for future prognosis. A few groups of researchers used Affymetrix microarray technology to study the gene expression profile of patients with CAN and identified differentially expressed genes related to CAN. It is of our interest to combine and compare the results from these different studies and find genes commonly predictive of CAN. Microarray experiment data from two studies were obtained upon requests. A discrimination method was applied on the gene expression measurements and used to combine across studies. An informative set of genes were found to be highly predictive of the presence of CAN and more discussions will be presented.

Selection and Validation of Normalization Methods for c-DNA Microarrays Using Within-Array Replications

❖ Yue Niu, Princeton University, Dept of Operations Research and, Financial Engineering, Princeton, NJ 08544, *yniu@princeton.edu*; Jianqing Fan, Princeton University

Key Words: normalization, validation test, empirical Bayesian estimator, genewise variance, within-array replications, cDNA microarray

Normalization of microarray data is essential for multiple-array comparisons. Several normalization protocols have been proposed based on different biological or statistical assumptions. A fundamental problem arises whether they have effectively normalized arrays. In addition, for a given array, the question arises naturally how to choose a method to most effectively normalize the microarray data. We propose a method to compare different normalization techniques. The test statistics involve estimating the genewise variances. This is accomplished by using several novel methods, including empirical Bayes methods for moderating the genewise variances between genes and the smoothing methods for aggregating variance information with similar intensity levels. And the P-values are estimated based on a normal or chi-square approximation. The method is illustrated by simulation and real data.

Inverse Regression Estimation for Censored Data

❖ Nivedita Nadkarni, University of Wisconsin-Madison, 1521 E Franklin Street, Apt C 102, Chapel Hill, NC 27514, *niveditan@gmail.com*; Michael Kosorok, The University of North Carolina at Chapel Hill

Key Words: right censored data, sufficient dimension reduction, inverse probability of censored weighting, inverse regression

An objective of analyzing survival data via regression is to develop a predictive model given covariates. An important step in formulating the model involves variable selection. Selection of the influential predictors is critical and becomes complicated if the data has high dimensional covariates, as is often the case in many clinical trials and more recently microarray studies. In addition to selection, assessment of predictor performance is also crucial. In this paper, we develop sufficient dimension reduction methodology via inverse regression for censored data. An inverse probability of censoring weighted approach is implemented to facilitate variable selection. A simulation study and several data analyses demonstrate the effectiveness of the proposed procedure.

Microarray Analysis: P-Values, Filtering, and Multiple Testing Adjustments

❖Ann Hess, Colorado State University, Campus Delivery 1877, Fort Collins, CO 80523, *hess@stat.colostate.edu*

Key Words: microarrays, multiple testing, filtering

Microarrays allow scientists to examine the expression of thousands of genes simultaneously. With one or more hypotheses being tested for each gene, a multiple testing adjustment is certainly warranted. However, in some cases the distribution of p-values and large number of tests mean that only a few genes are found to be significant after applying a multiple testing adjustment. Filtering allows us to reduce the number of tests and possibly allow more genes to pass the multiple testing criteria. Here we provide examples of real p-values distributions resulting from microarray experiments and examine various criteria used for filtering.

Distribution-Based Classification with Application to Gene Expression Data Analysis

❖Xuelian Wei, University of California, Los Angeles, Department of Statistics, 8125 Math Sciences Bldg., Los Angeles, CA 90095-1554, *wxl@stat.ucla.edu*; Ker-Chau Li, University of California, Los Angeles

Key Words: Classification, Prediction, Machine learning, Cancer classfication, Microarray, Gene expression data

Distribution-based classification (DBC) is a new technique for class prediction. Unlike univariate classification methods, such as gene voting, which treat each variable separately, DBC considers the joint effects of all variables simultaneously. Unlike K-nearest neighbor method (KNN) or Support vector machine (SVM), which utilize only a small proportion of training samples (the K nearest neighbors or the support vectors), DBC exploits the joint prediction strengths of all training samples. We present a geometric configuration to illustrate when our method have a clear advantage over other methods. When applying to 22 gene expression datasets for cancer classification, DBC performs as efficiently as the SVM and naive Bayesian. All outperform other learning methods (KNN, linear discriminant analysis, and decision tree).

197 Mixtures Models and Their Applications ● ✪

Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Mixture or Homogeneous: Modeling Placebo Response

❖Thaddeus Tarpey, Wright State University, 120 MM Building, Dayton, OH 45435, *thaddeus.tarpey@wright.edu*; Dong Yun, Wright State University; Eva Petkova, New York University

Key Words: EM Algorithm, skew normal

A common problem in statistical modeling is to distinguish between finite mixture distribution and a homogeneous non-mixture distribution. Often mixtures of normal densities are indistinguishable from homogenous nonnormal densities. We introduce a population-based EM algorithm for mixtures to illustrate what happens when a mixture is fitted to a homogeneous non-mixture distribution. An example regarding the nature of a placebo response in drug treated depressed subjects is used to illustrate ideas.

Mis-specification Bias in Poisson Mixture Models: Gamma versus Lognormal

❖Woollcott Smith, Temple University, Box 86, West Tisbury, MA 02575, *wksmith@temple.edu*; Shuo Zheng, Temple University

Key Words: Negative binomial, generalized linear models, mixture models

The Poisson-gamma mixture model is often used in modeling over-dispersed count data. This model is computationally tractable and is included in many generalized linear model packages. In many cases a more natural model for the variation in the random Poisson mean is the lognormal distribution. Since the mean and variance functions for both models are identical, mis-specification of the model can have little impact on statistical inference. However, for certain model properties, for example, the probability of nonzero counts, misspecification bias can exceed 50%. The implications of these results for bioassay and environmental monitoring are explored. Also, an efficient numerical integration method for the Poisson-lognormal is presented.

On Some Computational Issues in Marginal Latent Mixture Analysis

❖Yan Yang, Arizona State University, Department of Mathematics and Statistics, Tempe, AZ 85287, *yy@math.asu.edu*; Douglas Simpson, University of Illinois

Key Words: Correlated data, EM, Generalized estimating equation, Quasi-Newton, Zero-inflated

Data with bound-inflated responses are common in many areas of application. Often the data are bounded below by a real number (e.g., zero) with excess observations at the boundary value. We consider a general class of latent mixture models for inflated discrete and semi-continuous data that combines a degenerate distribution at the bound and a discrete or censored distribution. The latency resulting from not being able to identify which distribution has generated a boundary value leads to a pseudo-likelihood for correlated bounded data that cannot be factorized. We implement both the EM and Quasi-Newton algorithms to estimate the class of mixture models and compare the two methods. The asymptotic covariance matrix is adjusted by the sandwich estimator using the theory of generalized estimating equations. The methods are illustrated with an ultrasound safety study in laboratory animals.

Parametric Mixture Model for Survival Data

❖Ying Zhang, Wyeth, Collegeville, PA 19426, *zhangy8@wyeth.com*; Jagbir Singh, Temple University

Key Words: Parametric mixture model, EM algorithm, Maximum Likelihood Estimation, Model selection

The analysis of failure-time data is usually complicated by the presence of censoring so that the regular estimation methods need to be modified. In this talk, we propose a parametric mixture model and its estimation based on maximum likelihood. An Expectation-Maximization (EM) algorithm is implemented to achieve the maximum likelihood estimation of the parametric mixture model for survival data. Furthermore, we develop a statistic based on Bayesian Information Criterion (BIC) for model selection. The parametric mixture model is tested on both simulated data and real data. It is seen to work reasonably well, provides great flexibility and easy to interpret. Since the maximum likelihood method is used to estimate the parametric mixture model, the statistical inference is not difficult to follow.

Prediction by Trajectory Modeling Approach

❖Nami Maruyama, Kitasato University, 5 9 1 Shirogane Minatoku, Tokyo, 1088641 Japan, *nami.maruyama@pfizer.com*; Fumiaki Takahashi, Kitasato University; Hajime Uno, Kitasato University; Masahiro Takeuchi, Kitasato University

Key Words: mixture distribution model, latent class, EM algorithm, prediction, predictive model

In longitudinal data, the interest often lies in the repeatedly measured variable itself. However, in some situations the changing pattern of the variable over time may contain information about a separate outcome variable. In such a situation, longitudinal data provide the opportunity to develop predictive models of future observations of the separate outcome variable given current data for an individual. In particular, longitudinally changing patterns of repeated measurements of a variable, or trajectories, measured up to time t can be used to predict an outcome measure or event that occurs after time t. We propose a predictive model based on latent classes of trajectories, which is fit using EM algorithm, and show how to get model estimates with other covariates included in the model. Applications of our methodology are demonstrated through an example of a longitudinal study.

Mutual Information for the Mixture of Two Multivariate Distributions

❖Walfredo Javier, Southern University-Baton Rouge, Baton Rouge, LA 70813, *waljavi9@aol.com*; Arjun K. Gupta, Bowling Green State University

Key Words: mutual information, mixture of two normal distributions, normal linear and quadratic forms

Mutual Information for a multivariate random vector is a measure of dependence among the component random vectors; it is zero when the components are independent, positive otherwise. The paper derives the mutual information for a mixture of two multivariate normal distributions without the assumption that the two component normal distributions be widely separated. This is accomplished by partitioning the sample space into two disjoint subspaces in each of which the series expansion for ln(1+y) is valid.

198 Missing Data in Clinical Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Handling Missing Data in Long-Term Clinical Trials

❖Wei Deng, Amylin Pharmaceuticals, Inc., 9360 Towne Centre Drive, San Diego, CA 92121, *wei.deng@amylin.com*; Xuesong Guan, Amylin Pharmaceuticals, Inc; Chien-Feng Chen, Amylin Pharmaceuticals, Inc.

Key Words: longitudinal trials, last observation carried forward, missing data, mixed model repeated measure, multiple imputation, conditional model

This talk was motivated by a completed long-term clinical trial. The primary outcome was observed at different visiting schedules for different patients over three years. Our main interest is to estimate the change in primary outcome from baseline at the end of the trial. About one third patients withdrew from the trial prematurely. The impacts of missing data on inferences depend on the missingness process. Ignoring the missingness process may induce bias and inefficiency in the estimates. Thus the underlying missingness mechanism is explored and a few statistical models are considered, including last observation carried forward (LOCF), mixed model repeated measure (MMRM), multiple imputation (MI) and conditional model. A simulation study similar to the trial is also performed to compare these models.

Something Might Be Missed in Missing Data Analysis

❖Xiang Guo, sanofi-aventis, 35 River Dive South, APT1904, Jersey City, NJ 07310, *xiang.guo@sanofi-aventis.com*; Guangrui Zhu, Eisai Medical Research Inc.

Key Words: Fictitious assumption, Minimum Homogeneous Group, Missing data, Randomization

To make inference on missing data, an essential step is to posit assumptions about missing values. Those assumptions can dramatically affect the analysis result. It is hard to tell which result is unbiased since the methods are based on unverifiable assumptions. This core difficulty is generally not highlighted by the modelers, and the results often be mistakenly weighted by decision makers. We introduce a concept of fictitious assumption and distinguish it from the traditional concept of statistical modeling assumptions where observed data can verify the assumptions to certain degrees. Viewing the missing data handling methods from fictitious inference point of view can make the dependency of the inferences transparent to the decision makers. For available missing data methods, we use fictitious formulation to demonstrate the impact of the fictitious assumptions on the inference.

A Comparison of Imputation Methods in a Randomized Active Comparator Pain Relief Clinical Trial

❖Adam Hamm, Rho, Inc., 6330 Quadrangle Drive, Chapel Hill, NC 27517, *adam_hamm@rhoworld.com*

Key Words: Missing Data, BLUPs, Imputation

We evaluate the properties of a mixed model analysis with various imputation methods for handling missing data in a randomized, double-blind active comparator Phase III trial evaluating the efficacy of a pain relief treatment. We discuss a method for generating a dataset with complete outcome measures over time based on the results of an analysis of available blinded data, and then use this complete dataset to compare the various methods of imputation and determine the bias of each method under different assumptions for possible treatment effect. The baseline observation carried forward (BOCF) and last observation carried forward (LOCF) imputation methods are contrasted with an analysis of the available data (AVAL), an analysis that imputes missing data via a method of generating empirical best linear unbiased predictors (BLUP), and an analysis of the simulated complete dataset.

Some Remarks on Multiple Imputations in Longitudinal Data Context

❖Mohammed K. Alam, Kendle International Inc., 1200 Carew Tower, 441 Vine Street, Cincinnati, OH 45202, *mkalam27@yahoo.com*; Marepalli B. Rao, University of Cincinnati; Ramesh N. Amatya, Kendle International Inc.; Claudia Lara, Kendle International Inc.

Key Words: Multiple Imputations, Longitudinal Data, Clinical Trials, Missing Data, Inference, Dropout

The focus is on Longitudinal Data Collection in Clinical Trials in which each patient in the sample is observed at some specified n time points with respect to a well-defined response measure. For each patient in the sample, either we have a complete record of the data at all time points of interest or some data are missing at some time points. This scenario includes dropouts, who drop out of the study completely after a certain time point. Molenberghs et al (2004) have focused on analysis of data consisting of complete records and dropouts. Our scenario is more general. There are two predominant reasons in imputing missing data points in our scenario. One is to have a complete record of responses for each patient in the study. Other reason is to examine the effect of imputation on inference. Some of the issues in this context will be discussed.

A Resampling-Based Approach To Test If a Model Provides Unbiased Parameter Estimates Under Various Missing Data Assumptions

❖Maria M. Sanchez, Merck & Co., Inc., RY34 A316, PO Box 2000, Rahway, NJ 07065, *matilde_sanchez@merck.com*; Edmund Luo, Merck & Co., Inc.

Key Words: Resampling Method

For the three well defined missing data mechanisms (i.e., MCAR, MAR and NMAR), unbiased parameter estimates could be obtained under MCAR and MAR as long as the underlying model accommodates the missing data assumptions. While it is impossible to test the assumptions of missing data, we propose a resampling based approach to test if the model accommodates the missing data assumptions via checking if it provides unbiased parameter estimates. The same technique can also be applied to test if a model provides unbiased parameter estimates by modeling the missing data mechanism for NMAR.

Properties of Missing Data Imputation Methods: Baseline or Worst Observation Carry Forward

❖Jun Shao, University of Wisconsin-Madison, 1300 University Ave, Madison, WI 53706, *shao@stat.wisc.edu*; David Jordan, Abbott Laboratories; Yannis Jemiai, Cytel Inc.; Yili Pritchett, Abbott Laboratories

Key Words: baseline carry forward, bias, worst observations, clinical trials, multiple visits, dropout

Baseline-observation-carry-forward (BOCF) or worst-observation-carry-forward (WOCF) is sometimes applied to handle patient dropout in clinical trials with multiple scheduled visits. To many practical users, it is not clear what kind of effects we can estimate using data after applying BOCF or WOCF, and when and how statistical inference can be made using BOCF or WOCF data. We will address these issues in this presentation. After demonstrating the impact of BOCF or WOCF on study power via simulation results, we point out that sample means based on BOCF or WOCF data estimate a fracture of completer's sample mean where the multiplication factors are related to dropout rate, or the dropout rate and the percentage of worsening cases. We also demonstrate how to assess the variability of the sample means based on BOCF or WOCF data, which will ensure the accuracy of the statistical inferences.

Partial FEFI for Incomplete Tables with Covariates

❖Shin-Soo Kang, Iowa State University, 3308 Valley View, Ames, IA 50014, *sskang@iastate.edu*; Kenneth Koehler, Iowa State University; Michael D. Larsen, Iowa State University

Key Words: Complete Case Analysis, Fractional Imputation, Multiple Imputation, Wald Statistic

The information of covariates are available to do a partial fully efficient fractional imputation (FEFI). The new method, PFEFI with logistic regression is proposed to construct complete contingency tables. Jackknife method is used to get a standard errors of log-odds ratio from the completed table by the new method. Simulation results, when covariates have more information about categorical variables, reveal that the new method provides more efficient estimates of log-odds ratio than either multiple imputation (MI) based on data augmentation or complete case analysis.

199 Modeling Time-to-Event Data from Clinical Trials ● ✪

Biopharmaceutical Section

Monday, July 30, 2:00 pm–3:50 pm

Controlling for Time-Dependent Confounding Using Marginal Structural Models in the Case of a Continuous Treatment Covariate

❖Ouhong Wang, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, *owang@amgen.com*; Trevor McMullan, Validant Consulting

Key Words: Time Dependent Confounding, Marginal Structural Models, Inverse Probability of Treatment Weighting, Counterfactuals

Marginal Structural Models, whose parameters are estimated by Inverse Probability of Treatment Weighting (IPTW), can be used as an alterative to g-estimation to adjust for time dependent confounding, eliminating the need to construct the full counterfactural dataset. When the treatment variable is binary, the number of counterfacturals that need to be generated are 2**t per patient where t is the number of time points. When the treatment variable is continuous the number of counterfacturals needed is no longer finite. In these situations the performance of IPTW is not well characterized. Here we model time dependent confounding using IPTW and a continuous treatment variable with both observed data and bootstrapped simulated data. These results are compared to standard time dependent Cox proportional hazards models, as well as IPTW results with ordinal treatment.

Evaluating Treatment Effect in Randomized Trials with Learning Effect: An Application to Contraceptive Device Trials

❖Pai-Lien Chen, Family Health International, PO Box 13950, Research Triangle Park, NC 27709, *pchen@fhi.org*

Study participants in a randomized device trial must use study devices as designed to evaluate the true effect of devices. Nevertheless, the learning process of study devices is highly related to study outcomes and can be device-dependent. Without considering device-dependent learning effects, the conclusion of device trials will suffer potential bias. In this study we propose using a joint modeling approach to incorporate the learning effect into treatment effect evaluation. First, a generalized growth curve model is used to describe the leaning process, and then we use Cox's regression to model treatment effect for time to event outcomes with the integration of learning effect into the model. The parameters of the models will be estimated jointly by a semi-parametric method. A randomized contraceptive device trial will be used to illustrate the usefulness of the proposed method.

Robust Assessment of Survival Differences Using Multiple Testing Procedures

❖Larry Ma, Temple University, 129 Aileen Dr, Lansdale, PA 19446, *larry_ma@merck.com*; Devan V. Mehrotra, Merck & Co., Inc.; Woollcott Smith, Temple University

Key Words: Time-to-Event Data, Survival Difference, Weighted Log-Rank Statistics, Multiple Testing Procedures, Robustness

The log-rank test is optimal for assessing differences between two survival functions when the hazards are proportional to each other. However, if the proportional hazards assumption is not tenable, a weighted log-rank test is often used in which the pre-specified weight function assigns more weight to early or late events. Since the performance of a single weighted log-rank statistic is sensitive to the choice of the weight function, we suggest the simultaneous use of multiple weighted log-rank statistics when some prior information is available or there is no prior information at all. The equality of the two survival functions is tested using a pre-specified multiplicity adjustment. The MTPs that we consider include extensions of Fisher's and Simes' methods to the setting of correlated test statistics. We make recommendations based on simulation-based test size and power results.

An Empirical Test for Using a Cox Proportional Hazards Model with Time-Varying Hazard Ratios in the Analysis of Cardiac Events of Hypertensive Patients

❖D. Das Purkayastha, Novartis Pharmaceuticals, 44 Center Grove Road, Apt A 13, Randolph, NJ 07869, *das.purkayastha@novartis.com*; Lynn Ashcraft, Novartis Pharmaceuticals

Key Words: Time-varying hazard ratio, Cox proportional hazards model, Stationarity, Autocorrelation

A conventional Cox proportional hazards model postulates hazard ratios to be constant over time. If we consider time dependent covariates in the model with an assumption of time-varying hazard ratios as functions of weighted lag effects of change-point hazards, then the proportionality is disturbed in the model. In such models, in order to justify the clinical data and the model used for analysis, it is imperative to test for the stationarity of such a series of hazard ratios. This paper explains their stationarity conditions for their applications in the analyses of hazard models with an empirical exposition with 24 weeks of data of cardiac events from 200 hypertensive patients. Along with the empirical illustration, the paper also provides an analytical check of the applicability of a Cox model and stationarity conditions for hazard ratios with clinical data.

Nonproportional Hazards in Time-to-Event Analysis

❖Yanqiong Zhang, Merck & Co., Inc., 81 Shelley Circle, East Windsor, 08520, *yanqiong_zhang@merck.com*; Amarjot Kaur, Merck & Co., Inc.

Key Words: quantitative nonproportionality, qualitative nonproportionality, asymptotic relative efficiency

Cox proportional hazard (PH) model (1972) is the most commonly used method for time-to-event analysis and is referred to as a semiparametric model. Many a time in practice, the proportional hazard assumption does not always hold due to potentially two reasons. One reason is due to crossover of two hazard functions, where the hazard ratio could be >1 during one time period and <1 in the other period, also referred to as qualitative nonproportionality. The other reason of nonproportionality could arise where the hazard ratio doesn't keep constant over the time periods but no crossover of the hazard ratios occurs, also referred to as quantitative nonproportionality. An exploratory investigation of the power loss as well as asymptotic relative efficiency for the score test will be carried out in the presence of the quantitative nonproportionality.

The Impact of Proportional Hazards Assumption on the Late Onset Survival Data

❖Guoguang Ma, Amgen Inc., 1120 Veterans Blvd, ASF 21-2, South San Francisco, CA 94080, *gma@amgen.com*; Kathy Harris, Amgen Inc.

Key Words: survival analysis, proportional hazards assumption, Cox model, log-rank test, clinical trials

The log-rank and Cox tests are powerful under the assumption of proportional hazards that the relative risk remains constant over time. While these tests can be well powered to detect some differences between treatment groups that do not satisfy the assumption, they can have poor power to detect differences when cumulative incidence curves that are initially equal but later diverge or that initially diverge but later approach one another. This issue is usually raised in the analysis of clinical trials data when an investigational drug has a delayed treatment effect. In this work, we evaluate the impact of the proportional hazards assumption on the late onset survival data through simulation studies where we assume exponential distributions for the survival time and uniform distributions for the dropout time. The log-rank and Cox tests are also compared to the logistic regression model.

On Selection of Survival Models with Random Effects

❖Yanhua Zhang, University of California, Davis, 1121 Westbriar Dr., Apt 11, Richmond, VA 23238, *yhzhang@wald.ucdavis.edu*

Key Words: fence method, Framingham Heart Study, mixed effects, model selection, proportional hazards, survival analysis

We propose a procedure for selection of mixed effects models for survival analysis. The procedure is based on a recently developed method, called fence method (Jiang et al. 2006). The idea involves a procedure to isolate a subgroup of what are known as correct models (of which the optimal model is a member). While the current model selection procedures for survival analysis are mostly Bayesian, our procedure provides an alternative approach.In particular, we apply the method to proportional hazards models with random effects. The results of a Monte Carlo simulation and the analysis of the Framingham Heart Study data are reported.

200 Statistics for Customer Relations ● ✪

Business and Economics Statistics Section

Monday, July 30, 2:00 pm–3:50 pm

Scoring Customer Probability of Runoff for Retention

❖Huaiyu Ma, GE Global Research, One Research Circle, Niskayuna, NY 12309, *mah@ge.com*; Deniz Senturk-Doganaksoy, GE Global Research; Kareem Aggour, GE Global Research

Key Words: retention, runoff, survival analysis, regression, prediction

Effective customer retention is key for enabling a company to grow organically. Developing robust models with high accuracy on customer retention is difficult. A customer's past behavior as well as other factors such as customer interactions must be among the predictors for a model to be effective. Capturing these types of predictors correctly and making them actionable is not trivial, however. Multiple statistical techniques are used for capturing these metrics. Survival analysis is then used to create X's that capture the past customer behavior. Logistic regression is used to score customers and cluster analysis bucketizes them based on their likelihood to runoff prematurely. These scores provide a means to make retention actionable for the business marketing and sales teams. Sales and marketing campaigns have already been launched based on the results of this modeling work.

Customer Profiling and Mining with Multidimensional Statistical Analysis: The Case of Retail Banking Industry

❖Ming Xie, IBM China Research Laboratory, Building 19 Zhongguancun Software Park, 8 Dongbeiwang WestRoad Haidian District, Beijing, 100094 China, *xieming@cn.ibm.com*; Wen Jun Yin, IBM China Research Laboratory; Jin Dong, IBM China Research Laboratory

Key Words: Customer Profiling, Retail Banking, Banking Statistics

In retail banking industry, understanding exactly the characteristics and preferences of customers through statistical approaches is an effective

way to improve the performance of bank branches. In order to handle the complexity of data, in this paper we present multi-dimensional statistical analysis from diverse perspectives of banking customers, including demographic characteristics, transaction behavior and regional effects. Firstly, customer segmentation process is implemented by multi-dimensional clustering analysis. Statistical test procedures are then employed to mine the characteristics of various customer segments. Finally regression analysis is presented to forecast business trends of potential customers. The analysis is based on the data of one of the largest banks in China, and the results show the high potential of the study for practices in retail banking industry.

Assortment Planning Under Dynamic Customer Substitution and Cross-Selling

❖ Jin Dong, IBM China Research Laboratory, Building 19 Zhongguancun Software Park, 8 Dongbeiwang WestRoad Haidian District, Beijing, 100094 China, *dongjin@cn.ibm.com*; Xinxin Bai, IBM China Research Laboratory; Wen Jun Yin, IBM China Research Laboratory; Ming Xie, IBM China Research Laboratory

Key Words: Assortment Optimization, Dynamic Programming, Cross-selling, Substitution

Assortment planning is a process for retailers to decide which types of products to stock and the inventory level for each item. Demand substitution and cross-selling effect in customer service are the factors that influence the joint optimal choice. We formulate a single-period, state-dependent model to maximize the expected profit for retailers, in which customers dynamically purchase items with stochastic substitution and cross-selling behavior. The model is subject to given resources and service level constraints. Based on customer transaction databases, approaches are introduced to predict the initial demand and the possibilities for the next product to be chosen. Finally, numerical examples are presented to compare the optimal inventory strategies to those of traditional newsboy model and practical suggestion on assortment planning is then given.

Statistical Methods for Accounts Receivable Tracking and Forecasting

❖ Alejandro Veen, IBM T.J. Watson Research Center, 1101 Kitchawan Road, Route 134, Yorktown Heights, NY 10598, *aveen@us.ibm.com*; Yasuo Amemiya, IBM T.J. Watson Research Center

Key Words: Longitudinal forecasting, conditional logistic regression, event occurrence prediction, online tracking system

A business enterprise typically utilizes an IT system to keep track customer invoices and payments. To use the same system for projecting incoming cash flows and directing bill collection efforts, sound statistical analysis procedures need to be implemented. This talk presents such a procedure; in particular, we discuss a method for predicting the occurrence, timing, and amount of bill payment. As part of an online monitoring and tracking system, the method uses a moving time-window of data available in the system and refreshes the results regularly. A case study is presented to illustrate the usefulness of the method.

The Role of Six Sigma in Achieving Perfection

❖ Mohammad Quasem, Howard University, 2600th Street, Dept of ISDC School of Bus, Washington DC, DC 20059, *mquasem@howard.edu*; Daniel Owunwanne, Howard University

Key Words: Six, Sigma, Achieving, Process, improvement, Productivity

Six Sigma is a highly disciplined process that helps to focus on developing and delivering almost 100% perfect products and services. This process produces less than 3.4 defects in per million opportunities. It is a business improvement technology which eliminates all defects. The achievement of six sigma signifies high performance, reliability and value to the customers. In this study, we will examine the role of six sigma in productivity improvement.

Estimating Taint in a Population of Insurance Claims

❖ Andrew Ostarello, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *aostarello@exponent.com*; Duane Steffey, Exponent, Inc.

Key Words: Insurance, Taint, Sampling, PPS Sampling

The taint of an insurance claim is the dollar amount the insurance company overpaid on the claim (owing to poor claim management, fraud, corruption, or other errors). During a claims audit, companies are often interested in estimating the total taint across all claims they've managed. Since it's not feasible to audit all claims, sampling methods and estimators need to be devised. We developed a series of R programs to simulate the bias and variability of ratio estimators under three different sampling schemes: Simple Random Sampling, Stratified Random Sampling, and Probability Proportional to Size Sampling. We also varied the simulations across a variety of sample sizes, regions, and assumed taint distributions. We selected an estimator that minimized the variability and bias of the estimate of the total population taint.

Independence of Duplicate and Invalid Signatures on a Petition

❖ Mary Whiteside, University of Texas at Arlington, Box 19437, Arlington, TX 76019, *whiteside@uta.edu*; Mark Eakin, University of Texas at Arlington

Key Words: Nonlinear, Goodman, Petition signatures

This paper investigates the assumption that replicated signatures and signatures that are invalid for other reasons are independent. This would mean that the proportion of replicated signatures is the same for valid and invalid signatures. In this case, a non-linear estimate for the total number of valid signatures on a petition has been shown to be superior to the traditional Goodman estimate.

201 Analysis of High-Dimensional Data ●

ENAR, Section on Nonparametric Statistics, Biometrics Section

Monday, July 30, 2:00 pm–3:50 pm

Modeling Progression of Cerebrovascular Disease with Longitudinal MRI Data

❖ Qian Weng, University of California, Davis, 1072 Tanland Drive Apt 208, Palo Alto, CA 94303, *qweng@ucdavis.edu*

Key Words: transition model, neuroimaging, spatial correlation

Cerebrovascular disease (CVD) is associated with increased risk of cognitive decline. Magnetic resonance imaging (MRI) enables researchers to visualize abnormalities of cerebral white matter (WMH), a marker for CVD. We focus on the development of models for the progression of CVD via WMH as seen on sequential MRI. Data from a single MRI consist of measurements for several hundred thousand spatially defined voxels, each corresponding to a small volume of the brain. We assume we know the WMH status for each voxel and that biologically this damage is not repaired, but that errors may occur in the observed WMH data. We will present a model for analyzing these data, theoretical properties of the estimates, the small sample properties of the estimates based on simulations, and an application to data from sequential MRI obtained in a large scale longitudinal imaging study.

Quantile Rank Score Tests for Longitudinal Data

❖Huixia Wang, North Carolina State University, 2501 Founders Drive, Department of Statistics, Raleigh, NC 27695, *wang@stat.ncsu.edu*

Key Words: Information-sharing, Longitudinal data, Quantile regression

In this talk, I will introduce a quantile inferential method for analyzing longitudinal data. The developed rank score test is easy to implement and it is robust in performance. The proposed approach is applied to probe level microarray data for assessing differential gene expressions. I will focus on an enhanced quantile rank score test, which aims to improve the efficiency of the quantile rank score test at small samples through borrowing information across genes. The extension of rank score test to some other areas will be discussed.

Imaging Biomarkers and Sample Size Estimation

❖Chunming Li, Pfizer Inc., 2 Continental Court, South River, NJ 08882, *chunming.m.li@pfizer.com*

Key Words: magnetic resonance imaging, measurement error, log-normal, effect size, maximum likelihood

In the presentation, I will begin with a brief discussion of some common medical imaging modalities in clinical studies. I will elaborate on several study objectives using imaging biomarkers. However, the main focus will be on an imaging biomarker application and related study design and statistical methodology. I will discuss the statistical modeling, associated power calculation and simulation results. In the end, I will share some of the challenging issues in imaging clinical trials.

Use of Satterthwaite Degree of Freedom in t-Statistic for Microarray Experiments with Fewer Replications

❖Yu Ding, Merck & Co., Inc., 150 Oakland Avenue, D304, Lansdale, PA 19446, *yu.ding@merck.com*; Damaraju Raghavarao, Temple University

Key Words: Fewer Replications, Microarray Experiments, Satterthwaite Degree of Freedom

Many statistical methods to identify differentially expressed genes are based on a large number of arrays. However, many microarray experiments have fewer replications because of scarcely available resource. Most standard methods can not be used or are not accurate with small degrees of freedom for the tests. In this presentation, we use data from doubtfully expressed genes and unexpressed genes to construct a t-test, which has larger degree of freedom to identify differentially expressed genes for both direct and indirect design. We also apply our new methodology to ApoAI data. The results are promising.

Wavelet-Based Modeling of Clinical Outcomes Using Diffusion Tensor Image Data

❖William Prucka, The University of Alabama at Birmingham, 2013 Highlands Dr, Hoover, AL 35244, *prucka@uab.edu*; Christopher S. Coffey, The University of Alabama at Birmingham; Gary Cutter, The University of Alabama at Birmingham; Daniel S. Reich, Johns Hopkins University

Key Words: MRI, DTI, diffusion tensor imaging, wavelets, functional data analysis, multiple sclerosis

Diffusion tensor imaging (DTI) is an advanced MRI technique capable of in vivo characterization of the spatial and angular dependence of free water diffusion in tissue. Microstructural tissue components act as physical barriers, preferentially inhibiting diffusion along certain directions. This is evident in fibrous tissues, which exhibit greater diffusion parallel versus perpendicular to the fiber orientation. The measured anisotropy reveals fiber tracts, indicating pathways of connectivity and pathological disease processes. We present an exploratory wavelet-based spatial model for predicting outcomes using DTI data. The model is fit on wavelet transformed DTI data and the computed effects are back projected into the spatial domain using the inverse wavelet transformation. The model is tested via simulation and applied to a case/control study in multiple sclerosis.

Local False Discovery Rate Estimation Based on Bootstrap Null Distribution of Control Group Samples

❖Aixiang Jiang, Vanderbilt University, 571 Preston Research Building 6848, Nashville, TN 37232, *aixiang.jiang@vanderbilt.edu*; Yu Shyr, Vanderbilt University

Key Words: False discovery rate, local false discovery rate, Bootstrap, microarray

Many false discovery rate estimation methods often estimated the proportion of genes or features that truly came from the same distribution regardless of disease or control groups, but we have no idea about how accurate the proportion estimation was. To avoid this problem, our new method used only control group data to build up Bootstrap statistic null distribution, and then estimated fdr similarly to the Efron local fdr approach (Efron, 2004). This new method achieved the closest fdr estimations to true fdr values among five different fdr estimation methods in 72 simulation data sets with different settings. These five different local fdr estimation methods were then applied to Golub microarray data sets (Golub et al.,1999). All of these five fdr methods performed well in feature selection for classification and our new method was one of the best two methods.

Longitudinal Image Analysis of Change in Tumor/Brain Vascular Permeability Induced by Radiation

❖Xiaoxi Zhang, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, *xiaoxi@umich.edu*; Timothy D. Johnson, University of Michigan; Roderick J. Little, University of Michigan

Key Words: Repeated Measures, Image Analysis, Markov Random Fields, Reversible Jump Markov Chain Monte Carlo

It is known that the impotency of chemotherapy in curing brain cancer is largely due to the tight endothelia junction, which prohibits the passing of large chemotherapeutic molecules. It is hypothesized that the drug delivery of chemotherapy can be more efficient by differentially increasing vascular permeability in the tumor area relative to the normal brain through local radiation. In this work, we investigate the change in tumor/brain vascular permeability induced by radiation, which helps determining the optimal time of chemotherapy. The analysis is based on an edge preserving smoothing we developed earlier, under the assumption that we only observe a blurred and noisy observation of the "true" scene (i.e. the underlying vascular permeability) in an imperfect world. We model the "true" scene with a Markov Random Field and apply Reversible Jump Markov chain Monte Carlo algorithm.

202 Bayesian Analysis in Clinical Trials ●

Section on Bayesian Statistical Science, Biopharmaceutical Section

Monday, July 30, 2:00 pm–3:50 pm

Assessment of Bayesian Estimates of Biomarker's Surrogacy for a Time-to-Event Clinical Endpoint in a Single Trial

❖Qian Shi, The University of Iowa, Biostatistics, 525 Hawkeye Court, Iowa City, IA 52246, *qian-shi-1@uiowa.edu*; Mary K. Cowles, The University of Iowa

Key Words: adjusted association, Bayesian statistics, relative effect, surrogate endpoint, time-to-event data

The Relative effect (RE) and adjusted association (AA) proposed by Buyse are widely accepted measures of surrogacy. Cowles developed Bayesian joint models (BJM) to estimate these measures when the clinical endpoint of interest is time-to-event in a single trial. RE is intended to predict the treatment effect on the clinical endpoint based on its effect on the surrogate. AA measures the association between the clinical and surrogate endpoints, adjusted for treatment. Simulation studies under the BJM show that AA can be precisely and accurately estimated. Inclusion of an endpoint with high AA improves the prediction of censored failure times significantly. In contrast, estimation of RE has tremendous variability, and is influenced by the significance of the treatment effects. Alternatives to RE, multiple surrogate endpoints, and appropriate uses of a marker with high AA are considered.

A Bayesian Model for Sensitivity Analysis on Intent-to-Treat Population

❖ Niko Kacirorti, University of Michigan, 1419 Morton Ave, Ann Arbor, MI 48104, *nicola@umich.edu*

Key Words: Nonignorable Dropout Mechanism, Pattern-Mixture Models, Randomized Clinical Trials, Hypertension

The Trial Of Preventing Hypertension study is a randomized trial conducted to examine whether early pharmacological treatment in subjects with high blood pressure prevent or delay the development of hypertension. The primary study endpoint is the development of clinical hypertension over a 4-year period. During the study, some of the subjects dropped out prematurely. Commonly, analysis for such type of data is based on the intention-to-treat (ITT) principle where assumptions about the dropout mechanism are made. These assumptions are usually not testable using the observed data, and may introduce bias. This paper presents a method for performing sensitivity analysis under a Bayesian approach for different scenarios of the dropout mechanism. The model introduces an intuitive, easy-to-understand parameterization for performing sensitivity analysis, which is applied to the randomized trial.

Bayesian Subgroup Analysis in Clinical Trials

❖ Claudia Pedroza, The University of Texas Health Science Center at Houston, 1200 Herman Pressler Dr, RAS E831, Houston, TX 77035, *claudia.pedroza@uth.tmc.edu*

Key Words: bayesian analysis, subgroup analysis, clinical trials, center effects, multiple comparisons

Clinical trials test for an overall treatment effect in the target population. However, clinical investigators are often interested in estimates of treatment effect for various subgroups. Typically these trials are not designed nor powered to conduct subgroup comparisons and this type of analysis raises the concern of multiplicity effects. Here, we review the common practices in dealing with subgroup analysis in clinical trials. We then assess the performance of a Bayesian approach to subgroup analysis which uses hierarchical models with random effects. We use simulations to calculate the rate of false-positive/negative results under various scenarios. In particular, we investigate how well the Bayesian method performs in data sets with small total sample size as well as with large number of subgroups with small sample sizes.

A New Latent Cure Rate Marker Model for Survival Data

❖ Sung Duk Kim, University of Connecticut, Department of statistics, 215 Glenbrook Rd. U4120, Storrs, CT 06269, *sdkim@stat.uconn.edu*; Yingmei Xi, University of Connecticut; Ming-Hui Chen, University of Connecticut

Key Words: Classification, Cure rates markers, Latent variables, MCMC, Posterior distribution

We propose a new mixture model via latent cure rate markers for survival data with a cure fraction. In the proposed model, the latent cure rate markers are modeled via a multinomial logistic regression. The proposed model assumes that the patients may be classified into several risk groups based on their cure fractions. Based on the nature of the proposed model, a posterior predictive algorithm is also developed to classify patients into different risk groups. The proposed model not only bears more biological meaning, but also fits the data much better than several existing cure rate models based on the popular LPML measure. In addition, we develop necessary theories of the proposed models and efficient MCMC algorithms for carrying out Bayesian computation. A real data set from a prostate cancer clinical trial is analyzed in detail to further demonstrate the proposed methodology.

Bayesian Analysis of Incomplete Data in Crossover Trials

❖ Sanjib Basu, Northern Illinois University, Div of Statistics, DU 366, De Kalb, IL 60564, *basu@math.niu.edu*

Key Words: incomplete data, crossover trial, model comparison

Crossover designs are common in early phases of clinical trials and in bioavailability and bioequivalence studies. The observed data from a crossover trial can be incomplete due to dropout or other reasons. We propose a Bayesian joint model for the observed data and the missing process. We use Bayesian model comparison to compare different missingness models and consider application in data from a clinical trial.

Propensity Score Matching for Causal Inference When Multiply Imputing Missing Covariate Data

❖ Robin Mitra, Duke University, 115 Old Chemistry Building, Durham, NC 27705, *robin@stat.duke.edu*; Jerome P. Reiter, Duke University

Key Words: Propensity Score, Missing data, Multiple Imputation, Observational Study

Propensity Score Matching is used in observational studies to balance the distribution of measured covariates for treated and matched control units. This enables inference about the treatment effect to be based on comparable groups. When units have only partially observed covariates, propensity scores cannot be directly estimated. Multiply Imputing missing values will allow propensity scores to be estimated from the imputed data, however we may be sensitive to the accuracy of the imputation models. In this talk we present some ideas that reduce the dependence of propensity score matching on the choice of imputation models used and can be implemented together with standard Multiple Imputation Software for missing data.

Flexible Modeling of the Hazard Function for Breast Cancer Recurrence

Vanja Dukic, The University of Chicago; ❖ James Dignam, The University of Chicago, 5841 S Maryland Ave, MC 2007, Chicago, IL 60637, *jdignam@health.bsd.uchicago.edu*

Key Words: multiresolution models, Bayesian, survival analysis, hazard estimation, prognostic factor modeling, breast cancer

The hazard function may reveal important aspects of the failure process not apparent in the survival function. While nonparametric hazard estimates tend to be unstable, parametric models are typically too restrictive and semi-parametric approaches treat the hazard as a nuisance parameter, focusing instead on covariate effects. Here we examine the hazard of recurrence after treatment for breast cancer, with specific interest in a) the shape of the hazard and how disease characteristics influence it and b) whether there are one or more hazard changepoints. To address these questions, we extend the Bayesian multiresolution hazard (MRH) estimator (Bouman et al. 2007, 2005). The proposed hierarchical MRH model accommodates stratum heterogeneity and common covariate effects across strata. We contrast results with those from flexible parametric model extensions and nonparametric approaches.

203 Topics in Reliability ● ✪

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Monday, July 30, 2:00 pm–3:50 pm

On Risk Sets and Risk Factors

❖William D. Heavlin, Google, P O Box 2846, El Granada, CA 94018, *bheavlin@google.com*; Judith Koslov, Sun Microsystems, Inc.

Key Words: blocking, CART, Cox regression, Mantel-Haenszel test, mean cumulative functions, relative risk

Failure events associated with systems reliability admit the possibility of repair and continued availability. For analyzing systems reliability data, the principal approach uses mean cumulative functions (MCFs), the repairable-system analog to hazard functions. MCFs implicitly focus on the time-ordered slices at which failure events occur. Using models conditioned on these time slices (these risk sets), MCF-based methods can adapt the Mantel-Haenszel and Cox proportional hazard approaches for quantifying the relative-risk coefficients associated with particular factors. Further, when the number of candidate risk factors is large, a modification to classification trees ("CART") is suggested for model reduction. This modification is conditional on the risk sets, and exploits random permutations within the risk sets to calculate significance. We illustrate with a series of examples.

Fuzzy Estimation of Parameters in Statistical Models

❖Ranee Thiagarajah, Illinois State University, Dept of Mathematics, Campus Box 4520, Normal, IL 61790, *kthiaga@ilstu.edu*

Key Words: Fuzzy Log-Normal, Fuzzy Inverse Gaussian, Fuzzy Statistical Measures, Fuzzy Random Variable, Fuzzy Statistical Measures

We consider probability density function of some statistical models where the parameters are uncertain. The estimates of these uncertain parameters can be point estimates or interval estimates. Buckley (2004) introduced a method of estimation of parameters in parametric models, using a set of confidence intervals producing a triangular shaped fuzzy number for the estimator. In this paper, following Buckley (2004), we model the uncertainty in model parameters by using fuzzy number estimates for the parameters, and obtain the fuzzy probability density functions. This will be discussed for log-normal and inverse Gaussian distributions. The computation of several fuzzy statistical measures such as fuzzy mean, variance, skewness, kurtosis, etc. of these distributions will be illustrated through numerical examples.

Accelerated Destructive Degradation Test Planning

❖Ying Shi, Iowa State University, 119 Snedecor Hall, Ames, IA 50010, *yshi@iastate.edu*; Luis Escobar, Louisiana State University; William Meeker, Iowa State University

Key Words: reliability, large sample approximate variance, optimum ADDT plan, compromise ADDT plan, Monte Carlo simulation

Accelerated Destructive Degradation Test (ADDT) is used to estimate the time at which a fixed percentage of product have a strength less than the critical degradation level when operating at normal temperature conditions. An ADDT plan specifies a set of test stress levels of temperature and evaluation time and the units' corresponding allocations to each test level. A class of degradation models is used with specified planning values for the parameters and plausible distribution for the model variability. This talk describes methods to find good ADDT plans. We show how to obtain an optimum plan to minimize the large sample approximate variance of the maximum likelihood (ML) estimate of a specified quantity. We then propose a more useful compromise plan. Also Monte Carlo simulations are used to evaluate ADDT plans. The methods are illustrated with an application for an adhesive bond.

A Comparison of Maximum Likelihood Estimation and Median Rank Regression for Weibull Estimation

❖Ulrike Genschel, Iowa State University, 326 Snedecor Hall, Department of Statistics, Ames, IA 50011, *ulrike@iastate.edu*; William Meeker, Iowa State University

Key Words: maximum likelihood estimation, median rank regression, Weibull distribution

We investigate Maximum Likelihood Estimation (MLE) and Median Rank Regression (MRR) for estimating parameters of a Weibull Distribution under Type I censoring. Through simulation, we compare both methods using the empirical mean and RMSE of estimates as performance criteria. The simulation experiment evaluates the effects of several factors, such as the Weibull shape parameter value, number of censoring times, fraction failing and expected number of failures before a given censoring time. Both MLE and MRR tend to overestimate the Weibull scale parameter when the number of expected failures is small, but the MLE procedure typically has smaller bias RMSE. The same behavior generally holds for shape parameter and quantile estimates. However, under certain conditions when the expected number of failures is small, the MRR procedure may yield quantile estimates with better accuracy.

Availability and Cost Monitoring in Datacenters Using Mean Cumulative Functions

❖David Trindade, Sun Microsystems, Inc., 6005 Assisi Court, San Jose, CA 95138-2316, *david.trindade@sun.com*; Swami Nathan, Sun Microsystems, Inc.

Key Words: repairable systems, mean cumulative function, availability, reliability, datacenters, MTBF

In datacenters the availability of servers and other equipment is routinely monitored by recording uptime to total time. Uptime is based on the ability of the applications, machines, or the entire facility to provide an acceptable level of service and is usually defined as a percentage over a time window, such as a week or a month, plotted in calendar time. The cumulative uptime percentage is also used. However, such summary statistics can be poor measures, since they do not account for time dependence of failures and multicensoring. Furthermore, since availability is a function of both outage frequency and duration, there is no distinction between large numbers of small outages and small numbers of large outages. This paper describes approaches based on extensions to Mean Cumulative Functions used successfully at Sun Microsystems to analyze the reliability of repairable systems.

Estimation from Aggregate Data

❖Marc Fredette, HEC Montreal, 3000 Cote Sainte Catherine, Montreal, QC h3t2a7, *marc.fredette@hec.ca*

Key Words: aggregate data, multi-state model

This work proposes a statistical methodology to handle aggregate data. We consider individuals moving among a finite set of states and observations are only made at different discrete time points. Aggregate data arise when the only information available at these time points is the total number of individuals in every state. In such a case, the history of each individual is lost and the time spent in a given state is unknown. Aggregate data arises

in many fields such as medical science, ecology, social science, reliability, etc. Using a simple data completion technique, we will show how the mean sojourn time can be approximated and we will study under which conditions such an approximation is adequate. Applications will be displayed using data from the literature and results will be compared with other approaches.

Comparison Sequential Testing for Reliability: Optimal Truncation of Short Tests

❖Yefim H. Michlin, Technion, Quality Assurance and Reliability, Haifa, 32000 Israel, *yefim@technion.ac.il*; Genady Grabarnik, IBM T.J. Watson Research Center; Elena Leshchenko, Technion

Key Words: Sequential testing, Comparison testing, MTBF ratio, Sample number

The study deals with simultaneous testing of two systems, one "basic" (b) and the other "new" (n) - with an exponential distribution of their times between failures. It is checked whether the mean TBFn/MTBFb ratio (discrimination ratio - DR) equals a given value, versus whether it is smaller than the latter. Results are reported for tests of this type with large DR, i.e. with small average sample numbers (ASN). Such tests serve for fast comparison of real-life systems at early stages of development. In this paper it is established that the distribution density of possible pairs of first and second type errors decreases steeply with decrease of the ASN and with test truncation. This gives rise to difficulties in optimizing the solution. An algorithm and methodology are presented, yielding an effective assessment of test optimality and planning in fullest compliance with requirements.

204 Design and Analysis of Experiments and Other Statistical Methods

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Monday, July 30, 2:00 pm–3:50 pm

Design and Analysis of Material Characterization Experiments with Few Runs per Day

❖William Guthrie, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8980, Gaithersburg, MD 20899-8980, *will.guthrie@nist.gov*; Kenneth W. Pratt, National Institute of Standards and Technology

Key Words: hierarchical models, experiment design, variance components, random effects, statistical software, Standard Reference Materials

The National Institute of Standards and Technology characterizes the properties of a wide range of Standard Reference Materials for use in instrument calibration or measurement assurance. When certifying material properties, assessment of homogeneity and identification of other variance components that affect the measurement processes used for certification are critical. Measuring duplicate samples from multiple units of a given material in a single day is not always feasible, however, and rules out the use of standard hierarchical designs. This talk describes a class of split-hierarchical designs that can be used when only two or three runs can be made each day and outlines some associated analysis methods. Computational tools for the construction of these designs and analysis of the data using open-source statistical software with a spreadsheet interface are also described.

Analysis of Optimization Experiments

❖James Delaney, Carnegie Mellon University, Baker Hall 132B, Pittsburgh, PA 15213-3890, *jdelaney@stat.cmu.edu*; Roshan Joseph, Georgia Institute of Technology

Key Words: Empirical Bayes method, Practical significance level, Shrinkage estimation, Variable selection

The typical practice for analyzing industrial experiments is to identify statistically significant effects with a 5% level of significance and then to optimize the model containing only those effects. In this article, we illustrate the danger in utilizing this approach. We propose methodology using the practical significance level, which is a quantity that a practitioner can easily specify. We also propose utilizing empirical Bayes estimation which accounts for the randomness in the observations. Interestingly, the mechanics of statistical testing can be viewed as an approximation to empirical Bayes estimation, but with a significance level in the range of 15–40%. We also establish the connections of our approach with a less known but intriguing technique proposed by Taguchi. A real example and simulations are used to demonstrate the advantages of the proposed methodology.

Using Minimal Dependent Subsets To Evaluate Supersaturated Designs

❖Arden Miller, University of Auckland, Private Bag 92019, Auckland, none New Zealand, *miller@stat.auckland.ac.nz*

Key Words: factorial designs, screening experiments

This talk will consider the use of supersaturated designs for screening experiments. A good screening design is one that can look at a set of candidate factors in a limited number of runs and reliably identify those that are active. Recently, a number of "supersaturated" designs have been proposed as a way of increasing the number of factors that can be investigated using a given number of runs. However, there are questions about the reliability of these designs. This talk will explain how the minimal dependent subsets of a supersaturated design can be used to evaluate how many active factors could possibly be detected by that design.

Randomization of Factors in Multiphase Experiments

❖Theodore Bailey, Iowa State University, Department of Statistics, 121 Snedecor Hall, Ames, IA 50011-1210, *tbbailey@iastate.edu*

Randomization of treatments to units is a fundamental principle in the conduct of experiments. The object of this paper is to discuss and illustrate how the choice of randomization affects methods of analysis an interpretation of multiphase comparative experiments.

Change Point Detection Using Nonparametric Estimation of Profiles in Small Sample Setting

❖Kathryn Prewitt, Arizona State University, Math and Statistics Dept, PO Box 871804, Tempe, AZ 85287, *kathryn.prewitt@asu.edu*; Shilpa Gupta, Arizona State University

Key Words: local linear, nonparametric, changepoint

We use local linear nonparametric curve estimation to estimate the profiles where the number of observations per profile is small. Integrated squared error is used to compare the curves and the joint distribution is determined by simulation under the normal error assumption. If time permits, comparison with parametric methods is discussed along with a data example.

A Comparison of Model Combining Methods

❖Lihua Chen, The University of Toledo, Mail Stop 942, Toledo, OH 43606, *lihua.chen@utoledo.edu*; Panayotis Giannakouros, University of Missouri-Kansas City

Key Words: Bayesian Model Averaging, Adaptive Regression by Mixing, generalized linear models, model combining

The present work compares a prediction-based model combining method to other model combining methods in generalized linear models. We implement Bayesian Model Averaging using a BIC approximation in a way that can accommodate interaction terms and therefore can be compared to prediction-based model combining across all parametric model settings in which prediction-based model combining is implemented. We also present results for several other weighting schemes for model combining.

Within-Sample Prediction of Future Failure Times Based on Type-II Censored Samples from the Birnbaum-Saunders Distribution

❖Kevin S. McCarter, Louisiana State University, 161 Agricultural Administration Bldg, Baton Rouge, LA 70803-5606, *mccarter@lsu.edu*

Key Words: Prediction, Birnbaum-Saunders, Censoring, Simulation, Type-II, Calibration

In this study we investigate procedures for the within-sample prediction of the largest failure time in Type-II censored samples from the Birnbaum-Saunders distribution. The first procedure we consider uses the parameter-estimated conditional distribution of the largest order statistic to construct the prediction intervals. We show that intervals constructed in this way do not achieve nominal confidence levels, with the actual coverage probabilities depending upon several factors including the sample size and the observed number of failure times. Logistic regression is used to quantify this dependence, and the resulting model is used to develop a calibrated interval prediction procedure with improved true coverage probability. Simulation results show that the calibrated intervals attain their nominal confidence levels at most of the simulation factor settings considered in the study.

205 Mixture Models and Expectation Maximization

Section on Statistical Computing, Business and Economics Statistics Section, IMS, Section on Statistical Graphics
Monday, July 30, 2:00 pm–3:50 pm

Nonparametric Transformation of the Data to Obtain Bias Reduction in Kernel Estimation of the Distribution Function of Nonstandard Mixtures

❖Ennis McCune, Stephen F. Austin State University, Box 13040 SFA, Department of Mathematics and Statistics, Nacogdoches, TX 75962, *mccuneed@sfasu.edu*; Sandra L. McCune, Stephen F. Austin State University

Key Words: kernel distribution function estimation, nonstandard mixtures, bias reduction

Nonstandard mixtures occur when a random variable behaves in a continuous manner except at a countable number of discrete mass points. Polansky (2005) introduced a biased kernel estimator of the distribution function of nonstandard mixtures. In this paper, a new estimator of the distribution function of nonstandard mixtures with less bias than Polansky's estimator is obtained by applying to Polansky's estimator a nonparametric data transformation bias-reduction technique introduced by Swanepoel and Van Graan (2005). Statistical properties of the new estimator are determined and presented.

Fitting Mixture Distributions Using Generalized Lambda Distributions (GLDs): Examples, Comparisons with Normal Mixtures, and Computational Considerations

❖Wei Ning, Bowling Green State University, Department of Math and Stat, Bowling Green, OH 43403, *wning@bgnet.bgsu.edu*; E. J. Dudewicz, Syracuse University

Key Words: Mixtures, GLD, Normal distributions

Mixture models were studied by Karl Pearson in 1894 when he fitted a mixture of two normal distributions to data consisting of measurements on the ratio of forehead to body length in 1000 crabs. Most work since that time has used mixtures of normal distributions. We consider a model for mixtures of Generalized Lambda Distributions (GLDs). Since the GLD can fit the normal well, whenever a mixture of normals will fit data well, so will a mixture of the same number of GLDs. However, the GLD is a much broader family, and can do well in cases where the normal cannot. We include examples and computational considerations.

Acceleration of the EM Reconstruction Algorithm for PET Images Using Squared Iterative Methods

❖Constantine E. Frangakis, Johns Hopkins University, 615 N. Wolfe Street, E 3642, Baltimore, MD 21205, *cfrangak@jhsph.edu*; Ravi Varadhan, Johns Hopkins University; Christophe Roland, University of Science and Technology at Lille

Key Words: EM algorithm, emission computed tomography, polynomial extrapolation methods, convergence acceleration, ordered-subsets EM, SQUAREM

Positron emission tomography (PET) is a useful medical technology for imaging the activity of an organ. Reconstruction based on statistical models, using maximum likelihood estimation, provides accurate images. However, the time required for reconstruction is prohibitively high for practical applications, due to the slow convergence of EM algorithm used in the reconstruction. Here we demonstrate the effectiveness of a recently developed class of numerical schemes, called the squared iterative methods (SQUAREM) (Varadhan and Roland, 2004; Roland and Varadhan, 2005), in accelerating the convergence of PET image reconstruction using simulations on a phantom with 128 x 128 pixels. We also compare the performance of SQUAREM to that of the ordered subsets EM (OS-EM) method (Hudson and Larkin, 1994), which is a widely-used, fast reconstruction algorithm.

Improving the Efficiency of the Monte Carlo EM Algorithm Using Squared Iterative Methods

❖Ravi Varadhan, Johns Hopkins University, 2024 E. Monument Street, suite 2-700, Baltimore, MD 21205, *rvaradhan@jhmi.edu*; Brian S. Caffo, Johns Hopkins Bloomberg School of Public Health; Wolfgang Jank, University of Maryland

Key Words: EM acceleration, mixed models, extrapolation methods, SQUAREM

Implementation of the Monte Carlo EM (MCEM) algorithm is challenging mainly due to the difficulty in monitoring its convergence. Fixed sample size MCEM algorithms can be highly inefficient, since large MC samples are required to obtain even moderately accurate (2-3 significant digits) estimates. Caffo et al. (2005) have recently proposed an ascent-based MCEM algorithm that solves these problems. Here we evaluate whether the efficiency of the ascent-based MCEM algorithm can be further improved using a new class of numerical schemes called squared iterative methods (SQUAREM), which have recently been proposed to accelerate the convergence of the EM algorithm (Varadhan and Roland, 2004). SQUAREM can be easily implemented as it only requires the basic EM step. We present the results of two simulation examples: maximum likelihood (ML) estimation in multivariate-t and logit-normal models.

Generalized t-Copula and Its Application on Biometric

❖Wenmei Huang, Michigan State University, Dept Statistics and Probability, East Lansing, MI 48824, *huangw11@msu.edu*; Sarat Dass, Michigan State University

Key Words: Copula functions, t-copulas, generalized densities, biometric authentication, biometric fusion

When forming a D-variate joint distribution from the marginals and a copula function, a common assumption made is that the marginal distributions should be continuous. The data fusion application requires marginal distributions that are generalized, that is, consisting of both discrete and continuous components. Multivariate joint distributions are formed based on coupling the D generalized marginals with a family of t-copulas. Identifiability of the parameters associated with the generalized multivariate distributions is established. An EM algorithm is developed for the parameter estimation and a strong consistency result is established for the parameter estimates. The result can also be applied for Gaussian copula. The newly developed methodology is used to model the distribution of genuine and impostor matching scores arising in biometric authentication.

Comparison of the Six Sigma and Lean Sigma on the IT Management Processes

❖Genady Grabarnik, IBM T.J. Watson Research Center, 19 Skyline Dr, Hawthorne, NY 10583, *genady@us.ibm.com*; Larisa Shwartz, IBM T.J. Watson Research Center

Key Words: six sigma, it process management

Managing IT is the practice of overseeing and running a company's information technology (IT) investments and systems. Efficient and cost effective IT management processes are essential for a smooth operation of the enterprises. We apply two approaches: six sigma and lean sigma to some of IT service management processes in order to establish their stability under different type of statistical perturbations.

206 Online Evaluations, Online Discussions, Podcasts, and Electronic Assessment Systems

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Monday, July 30, 2:00 pm–3:50 pm

From Paper-and-Ink to Online Evaluation: A New Computer Maple TA–Based

❖Li Wang, Student, 240 Westcourt Place, Waterloo, ON N2L2R7 Canada, *wangli1222@hotmail.com*; Yulia R. Gel, University of Waterloo; Sean Scott, University of Waterloo

Key Words: Online Education, Web-based Assessment, Distance Education, Self-learning, Undergraduate Statistics Education, Elearning

The move to computer based evaluation is a natural outcome of the increasing use of information technologies to enhance learning. Our goal is to develop a set of online tools for distance education, evaluation and self-learning of undergraduate courses in statistics using Maple TA software. The proposed tool includes an extensive bank of questions and randomizes the order of questions and answers. Moreover, a series of parameter-based questions is developed in such a way that parameters are able to take on random values within a defined range. The effect is that any question can generate a different instance for each student, and the feedback given can also be particular to each student's parameters, instead of a general solution. In this talk we present outcomes of utilizing this tool for teaching an undergraduate time series course at the University of Waterloo.

Group Discussions in an Online Statistics Course: The Role of the Instructor

❖Michelle Everson, The University of Minnesota, Dept of Ed Psych University of Minnesota, 206 Burton Hall 178 Pillsbury Drive SE, Minneapolis, MN 55455, *gaddy001@umn.edu*

Key Words: distance education, active learning, collaboration, curriculum design, instructor role

Using group discussion assignments in online courses can provide students with opportunities to talk about their understanding of different concepts and ideas, in addition to helping their peers learn important course material. Unfortunately, few published descriptions of online statistics courses provide details about how discussion assignments can or should be structured, nor do they outline the instructor's role in these assignments. In this presentation, the author will share her experiences using group discussion assignments in the online environment. The way in which these assignments have evolved over the years will be outlined, and guidelines will be presented about ways in which instructors can go about creating relevant discussion assignments, evaluating these assignments, and participating in these assignments.

The Benefits of Using a Course Disk To Aid in the Instruction of Statistics Courses

❖Jamis Perrett, University of Northern Colorado, Campus Box 124, McKee Hall 521, Greeley, CO 80639, *jamis.perrett@unco.edu*

Key Words: statistics education, online learning, educational technology

A course disk in either CD or DVD format can prove vital to an online or distance course in statistics, but can also prove beneficial for traditional on-campus courses. A typical course disk may include syllabus and course outline, lecture notes and lecture outlines, assignments, links to online data or articles, example SAS files, stat tables, solutions, video lectures, tutorials, etc. This informational presentation will demonstrate the potential of a statistics course disk, applications to online/distance learning courses as well as on-campus courses. It will demonstrate technological advancements of a course disk over previously-used methods. It will highlight the benefits of a course disk vs. other online approaches, especially in the area of online course delivery, and will give a brief overview of the equipment and software needed for creation of a successful course disk.

Electronic Assessment Systems for Introductory Statistics Courses I

❖John McKenzie, Babson College, MathScience Division, Babson Park, MA 02457-0310, *mckenzie@babson.edu*

Key Words: applied, homework, quiz, review exercises

This is the first of three contributed presentations on electronic assessment systems for introductory applied statistics courses. This presentation will offer an overview of such systems. It will summarize their advantages and disadvantages. The past, present, and future of electronic assessment systems will also be presented. The next two speakers will describe their experiences with dealing with these systems for quizzes, homework, and review.

Electronic Assessment Systems for Introductory Statistics Courses II

❖Patricia Humphrey, Georgia Southern University, Department of Mathematical Sciences, PO Box 8093, Statesboro, GA 30460-8093, *phumphre@georgiasouthern.edu*

Key Words: Introductory Statistics, Electronic Assessment, Homework Systems, Quizzes, Review Exercises

Electronic assessment systems are coming of age. Whether generic (implemented by the institution) or tailored to a specific text, they can be of great benefit to instructors and students, or cause extra work. This author will discuss her experiences with both types: the benefits and the pitfalls.

Electronic Assessment Systems for Introductory Statistics Courses III

❖William Rybolt, Babson College, Math Science Division, Babson Park, MA 02457-0310, *rybolt@babson.edu*

Key Words: applied, homework, quiz, review exercises, guidelines

This is the third of three contributed presentations on electronic assessment systems for introductory applied statistics courses. This presentation offers an overview of principles we have adopted to facilitate designing, building, and using electronic quizzes. These principles have evolved from our attempts to satisfy the goals of the instructors, the complaints of the students, and difficulties present in both the creation and the usage environments. The presentation concludes by discussing guidelines for others based on our experience.

Use of Podcast Technology in Statistics Education

❖Joni Nunnery, Louisiana State University, 36318 Page Drive, Information Systems & Decision Science, Denham Springs, LA 70706, *jnunner@lsu.edu*

Key Words: podcast, tablet technology, statistics education, recording software

With the advent of the internet, many technologies have emerged as tools for enhancing the educational experience. One such technology is the podcast, a pre-recorded audio and/or video file made available on the internet to be played on any computer or mobile device at anytime. This presentation will give a brief description of podcasting, academic uses, and the associated benefits and weaknesses. While the name implies the use of an iPOD, this presentation will show that podcasting requires only a computer (MAC or PC), recording software, a microphone, and the internet. Demonstrations, incorporating Tablet technology, will be applied to the area of statistics education. Podcast mini-lectures will be created covering topics on generating histograms in Excel, normal probability distributions and z-tables, z-tests, and regression estimation with interpretation of accompanying output.

207 Measurement Error in Survey Data Collection

Section on Survey Research Methods, Social Statistics Section

Monday, July 30, 2:00 pm–3:50 pm

Modeling Rotation Group Bias and Survey Errors in the Labor Force Survey of Statistics Netherlands

❖Jan Van den Brakel, Statistics Netherlands, Post Office Box 4481, Room Number R416, Heerlen, 6401 CZ Netherlands, *jbrl@cbs.nl*; Sabine Krieg, Statistics Netherlands

Key Words: rotating panels, structural time series models, small area estimation, Kalman filter

The Dutch Labor Force Survey (LFS) is a monthly survey, based on a rotating panel design. A major problem is that systematic differences occur between the subsequent waves due to panel effects, which is known as rotation group bias (RGB). Furthermore the monthly sample size is too small to produce reliable monthly estimates with the generalized regression estimator. Therefore, a multivariate structural time series model that accounts for the RGB and the autocorrelation between the different panels is applied to the monthly data of the LFS. With this model, estimates are obtained that are corrected for the RGB. Furthermore the model borrows strength from data observed in preceding periods, resulting in a remarkable reduction of the standard error compared with the generalized regression estimates. This results in a substantial increase of the accuracy of the monthly unemployment figures.

Face-to-Face Interviews with Children

❖Marek Fuchs, University of Kassel, Social Science Department, NoraPlatielStrasse 1, Kassel, 34109 Germany, *marek.fuchs@uni-kassel.de*

Key Words: Children, face-to-face interview, cognitive resources, interviewer style, response error, behavior coding

Increasingly, children are becoming respondents in survey interviews. While juveniles are considered to be reliable respondents for many topics and survey settings it is unclear to what extend younger respondents provide reliable information. In this paper we will report results from a study using video captures of 150 face-to-face interviews with children aged 9 through 11. The interviews have been coded using behavior codes on a question by question level which provides behavior-related indicators regarding the question answer process. In addition, standard tests of cognitive resources have been conducted. Also, parents were asked a set of questions identical to the children's questionnaire. Using the responses from the parents as gold standard, we are able to assess the impact of the children's cognitive resources on respondent behaviors and ultimately on response accuracy.

Survey Design for Studies of Measurement Error in Physical Activity Assessments

❖Nicholas Beyler, Iowa State University, CSSM, 204 Snedecor, Ames, IA 50036, *beylern@iastate.edu*; Sarah M. Nusser, Iowa State University; Wayne Fuller, Iowa State University; Alicia Carriquiry, Iowa State University

Key Words: Calibration, Measurement Error, Physical Activity, Energy Expenditure

It is widely accepted that energy expenditure (EE) data obtained from 24-hour physical activity (PA) recalls and objective activity sensing devices are subject to considerable measurement error, but limited information is available about the sources and form of the measurement errors. We consider the problem of developing a sample design for a PA survey that includes 24-hour recalls and measurements from an objective sensing device. Our goal is to investigate measurement error models and calibration functions for EE. We propose a measurement model for both types of EE measurements, derive method of moments estimators of model parameters, and use data from a pilot study (Calabro 2006) to calculate preliminary estimates. Finally, we consider the cost structure of the different measurements in determining appropriate sample designs.

Evaluating the Assumption of Uncorrelated Error Terms in a Mixed-Mode Study with Validation Data

❖Courtney Kennedy, University of Michigan, 426 Thompson St, Room 4050, Ann Arbor, MI 48104, *ckkenned@umich.edu*; Stanley Presser, University of Maryland

Key Words: measurement error, mixed-mode, validation study, IVR, CATI, web

Survey data users frequently invoke the assumption of uncorrelated error terms. Scant research, however, has addressed its validity in surveys of human populations. The few existent studies have yielded two dominant findings: correlated errors are 1) most common between items pertaining to a shared topic and 2) not well predicted by demographic variables. This paper uses data from a validation study featuring randomized mode assignment (IVR, CATI, Web) to bolster the literature on this topic in two ways. First, we determine whether the previous findings replicate in a contemporary survey, noting that the survey environment has evolved in numerous ways during the past 20 years. Second, prior studies relied exclusively on data from in-person interviews, but we are able to test for interactions between the newer modes of data collection and the nature and prevalence of correlated errors.

Measurement Error in Calibration: An Application to the National Resources Inventory

❖Emily Berg, Iowa State University, CSSM, 208A Snedecor, Ames, IA 50011, *emilyb@iastate.edu*; Jean Opsomer, Iowa State University; Wayne Fuller, Iowa State University

Key Words: calibration, measurement error, survey sampling

Properly accounting for measurement error in calibration problems can greatly improve estimates and predictions. New data collection protocols are being implemented for the National Resources Inventory (NRI), a longitudinal survey of nonfederal land in the US. As one of the primary objectives of the NRI is to estimate change over time, it is essential that changes in land use are not confounded with the effects of new measurement techniques. In calibrating the new protocols against standard methods, it is important to acknowledge sources of error inherent in both procedures. A special study was conducted as part of the 2005 NRI to estimate the impact of new protocols, in the presence of multiple sources of variation. We present a measurement error model for the data from this calibration study and discuss estimation of regression coefficients and variance components.

208 Web and Telephone Survey Methods ●

Section on Survey Research Methods, Social Statistics Section

Monday, July 30, 2:00 pm–3:50 pm

A Comparison Between a Traditional RDD Survey and a Telephone Survey with Cell-Phone-Only Populations

❖Sunghee Lee, University of California, Los Angeles, 10960 Wilshire Blvd, Suite 1550, Los Angeles, CA 90024, *slee9@ucla.edu*; David Grant, University of California, Los Angeles; J. Michael Brick, Westat; Sherm Edwards, Westat

Key Words: Coverage bias, Cell-phone only population, RDD surveys

This study examines potential coverage bias of traditional RDD surveys associated with exclusion of cell-phone only populations. Unlike previous studies where in-person survey data were divided based on reported telephone usage status to examine the bias, this study overcomes that limitation by using data from two telephone surveys. The first comes from the 2005 California Health Interview Survey (CHIS), a large-scale traditional RDD survey. The second comes from a pilot study conducted as a part of CHIS 2005. In the pilot, adults were sampled from households with cell-phone only and then administered the full CHIS interview on cell phones. Because the pilot data represent those excluded in RDD surveys, comparisons of estimates between the two surveys will allow us to assess the coverage bias. We will use estimates across a wide range of demographics and health characteristics.

A New Application of Adaptive Web Sampling Designs

❖Hong Xu, The Pennsylvania State University, 4048E Bannock Burn Place, Charlotte, NC 28211, *hxu@stat.psu.edu*; Steve K. Thompson, Simon Fraser University; James L. Rosenberger, The Pennsylvania State University

Key Words: adaptive web sampling, adaptive sampling, anti money laundering, network sampling

Sampling from rare, hidden and hard-to-access populations and creating estimation methods to obtain efficient information from the sampled data are core challenges in sampling theory and applications. This paper builds on Adaptive Web Sampling, proposed by Thompson(2006). These new sampling strategies extend previous adaptive network sampling methods to have more flexibility in controlling sample coverage. Here we propose a new application to investigate financial crimes such as money laundering, where the criminal transforms illegally derived funds into an apparently legal source. Financial institutions are developing anti-money-laundering strategies and training their associates to detect these fraudulent financial activities. We show designs that can efficiently assist in detecting highly risky accounts using wire transactions.

Time Series Analysis of Census Internet Response

❖Fred Highland, Lockheed Martin Corporation, 12226 Detour Rd Box 36, New Midway, MD 21775, *fred.highland@lmco.com*

Key Words: Census, Internet, Time Series

Collection of census data over the Internet promises to be more respondent friendly, accurate and cost effective than other methods. However, the characteristics of Internet response over time are not well understood limiting the ability to accurately model, predict and manage take-up. This work theorizes that census Internet response can be modeled as a time series of exponentially decreasing response probabilities associated with stimulating events. The events can be represented as statistical distributions that characterize the events and model response. The theory is applied to census Internet response data provided by Canada, the US and Australia to determine the common characteristics of response events. The approach provides a means to model Internet response patterns over time when calibrated to population and the survey methodology characteristics.

Assessing the Applicability of Self-Anchoring Scales in Web Surveys

❖Frederik Van Acker, Vrije Universiteit Brussel, Pleinlaan 2, Brussels, 1000 Belgium, *frederik.van.acker@vub.ac.be*; Peter Theuns, Vrije Universiteit Brussel

Key Words: web surveys, context effects, self-anchoring scales, non-response

Self-anchoring scales were first mentioned in the 60s as rating instruments for which the end anchors are defined by the respondent himself, based on his own assumptions, perceptions, goals and values. The uses of such scales are legion and they were found very useful to reduce measurement bias in cross-cultural research. In this study we investigate whether contextual influences can be lessened or eliminated by using self-anchored scales. For this purpose a similar experiment as the one by Couper et al (2004), in

which they manipulate images placed in a web survey, was conducted. We hypothesize that self anchoring scales can reduce contextual bias since defining the scale typically requires recollection of several, affectively opposing, exemplary situations. The second part of the study investigates the impact of self-anchoring scales on nonparticipation and drop-out in surveys.

Comparing Web Survey Samples of Schizophrenic and Bipolar Patients with Concurrent RDD and In-Person Samples

❖ LinChiat Chang, Opinion Research Corporation, 2133 Cornell St, Palo Alto, CA 94306, *linchiat.chang@opinionresearch.com*; Channing Stave, Pfizer Inc.; Corinne O'Brien, Pfizer Inc.; Fred Rappard, Pfizer Inc.; Jill Glathar, Opinion Research Corporation; Joseph Cronin, Opinion Research Corporation

Key Words: web survey, mode comparison, sample quality, mental illness, schizophrenia, bipolar disorder

Web panel samples of patients with mental disorders face multiple problems, including coverage error, self-selection bias, and physicians cannot validate the purported diagnosis. To evaluate these web samples, we administered the same measures of disability to schizophrenic and bipolar patients via three modes: web, telephone, and in-person. Samples for the web and in-person modes were drawn from existing web panels, while the telephone sample was recruited via RDD methodology. Among schizophrenic patients, the web sample yielded significantly higher estimates of disability, and contained higher proportions of young, married, highly educated, White male respondents with full time employment. Mode differences among bipolar patients were much less pronounced. The significant differences in disability between modes were often eliminated after controlling for demographic differences.

Beyond Demographics: Are 'Webographic' Questions Useful for Reducing the Selection Bias in Web Surveys?

❖ Matthias Schonlau, RAND Corporation, 4570 Fifth Ave, Suite 600, Pittsburgh, PA 15213, *matt@rand.org*; Arthur VanSoest, RAND Corporation; Arie Kapteyn, RAND Corporation

Key Words: web survey, propensity score, selection bias, webographic variables, lifestyle variables

The main concern about Web surveys is that inference is affected by non-random selection of web survey participants (selection bias). One approach to reduce selection bias is to use propensity scores and a parallel phone survey. This approach uses demographic and additional so-called webographic or lifestyle variables to balance observed differences between web survey respondents and phone survey respondents. Here we investigate some of the webographic questions used by Harris Interactive, a commercial company specializing in Web surveys. Our webographic questions include choice of activities such as reading, sports and traveling and perceptions about what would constitute a violation of privacy among others. We use data from an existing probability sample representative of the US 50+ population and their spouses (Health and Retirement Study) and a corresponding web survey.

When Standard Regression Techniques Fail: The Role of Propensity Score Methods

❖ Michael Posner, Villanova University, Dept of Math Sci, 800 Lancaster Ave, Villanova, PA 19085, *michael.posner@villanova.edu*; Arlene Ash, Boston University

Key Words: Propensity Scores, Selection Bias, Model Misspecification, Causal Inference

When can propensity score methods solve problems that standard covariate adjustment cannot? While some view propensity score methods primarily as a tool for dimension reduction of a complex covariate space (with all important covariates measured), the problem of biased estimates of treatment effect can arise with even a single covariate. We examine a simple model, conceptually and via simulations, to show that standard regression can appropriately handle differing covariate distributions between treatment groups, but when model misspecification is also present, estimates from standard regression are biased. We compare analyses using crude estimates (no adjustment), regression, and propensity score methods. When covariate distributions differ by treatment group and models are mis-specified, only the propensity score method provides an unbiased estimate of treatment effect.

209 Contributed Poster Presentations

Section on Statistics in Epidemiology, Section on Government Statistics, Section on Health Policy Statistics, Section on Statistics and the Environment, Section on Survey Research Methods, Social Statistics Section, Section on Statistical Computing

Monday, July 30, 2:00 pm–3:50 pm

Modeling Tilletia Indica Infection on Wheat Using Nonlinear Mixed Models with a Sharp Turning Point

❖ Zhongwen Tang, Kansas State University, 101 Dickens Hall, Manhattan, KS 66502, *zhongwent@yahoo.com*; Shie-Shien Yang, Kansas State University; Karen Garrett, Kansas State University; I. Sharma, Punjab Agriculture University; R. L. Bowden, U.S. Department of Agriculture

Key Words: Nonlinear mixed model, Turning point, Bayesian, Maximum Likelihood, SAS, Winbugs

The infection and reproduction of Karnal bunt pathogen, Tilletia Indica, depend on the amount of sporidium encountering the host plants. A data driven nonlinear mixed model with a sharp turning point is built to precisely describe the relationship between the concentration of Tilletia Indica sporidium and the infection success percentage. Bayesian and Maximum likelihood methods are used to analyze the data. Model equations and turning points are estimated. SAS and WinBUGS are used for computation.

Waiting Time Distributions of Sleep

❖ Anna Mnatsakanova, National Institute for Occupational Safety and Health, 1095 Willowdale Road, Morgantown, WV 26505, *fma8@cdc.gov*; James Slaven, National Institute for Occupational Safety and Health; John M. Violanti, State University of New York at Buffalo; Bryan J. Vila, Washington State University; Cecil M. Burchfiel, National Institute for Occupational Safety and Health; Michael Andrew, Centers for Disease Control and Prevention

Key Words: Sleep quality, waiting times, distribution analysis

Waiting time distribution is a novel way to look at the quality of sleep. To elucidate sleep, we have characterized sleep quality by calculating the distributions of time from a state of sleep until first disturbance (waiting time distributions) over periods of time for each individual, and analyzed these waiting time distributions. Quality of sleep of participants from a study of stress and fatigue in police officers was categorized as good or poor using accelerometer-derived sleep variables. Twelve subjects, six of each quality, were chosen to explore the differences between sleep qualities with respect to waiting time distributions. Waiting time distributions were calculated and non-parametric analyses were performed. Nonparametric analysis

showed that there were significant differences in median times between the two groups.

A Goodness-of-Fit Test of Logistic Regression Model

❖Ying Liu, Kansas State University, 1604 Roof Drive D30, Manhattan, KS 66502, *yliu999@ksu.edu*

Key Words: logistic regression, goodness-of-fit, rejection rate

Logistic regression model is widely used in many areas. Before we can make valid statistical inference based on a logistic regression model, we must test whether the model fit the data adequately. Pearson Chi-square test and Deviance test can not provide correct p-value with one or more continuous covariates. Hosmer and Lemeshow test is standard goodness-of fit tests. Other recent tests are Osius and Rojek's test and Stukel's test. In this article, we purpose to approximate the "true" model by a partition logistic regression model, since it includes the logistic regression model as a special case. This partition model is used to construct goodness-of -fit test of a logistic regression model. Our simulation results show that the proposed test controls type I error rate well and that it has higher rejection rate than the other known methods when the assumed model is not correct.

Categorization of Sleep Patterns with Derived Actigraph Variables

❖James Slaven, National Institute for Occupational Safety and Health, MS 4050, 1095 Willowdale Rd, Morgantown, WV 26501, *cto8@cdc.gov*; Michael Andrew, Centers for Disease Control and Prevention; John M. Violanti, State University of New York at Buffalo; Bryan J. Vila, Washington State University; Cecil M. Burchfiel, National Institute for Occupational Safety and Health

Key Words: actigraphy, sleep patterns, cluster analysis, discriminant analysis

Actigraphy is an increasingly popular method of analyzing sleep cycles. Most studies that use actigraphy select only a small set of sleep variables to analyze and report. These variables are useful as descriptive statistics and as a way to categorize sleep quality, but do not fully parameterize sleep patterns. Additional information can be derived from variables regarding a participant's in-bed activity and circadian rhythm, which can aid in sleep quality characterization. In a study of health outcomes with police officers, 228 participants had their sleep analyzed with actigraphy. Sleep was categorized as good or poor using basic sleep variables. Discriminant and cluster analyses were performed on the additional sleep variables available through actigraphy. These additional variables had low error rates when discriminating between sleep qualities.

Logistic Regression, Classification Tree, and Boosted Regression: A Comparison of Models in the Prediction of Pregnancy-Induced Hypertension

❖Paul Kolm, Christiana Care Health System, 131 Continental Drive, Suite 202, Newark, DE 19713, *pkolm@christianacare.org*; Deborah Ehrenthal, Christiana Care Health System; Claudine Jurkovitz, Christiana Care Health System; Matthew Hoffman, Christiana Care Health System; James Bowen, Christiana Care Health System

Key Words: Data mining, Logistic regression, Tree classification, Boosted regression

The development of data-mining methods, and software, have facilitated building regression models from large databases in terms of number of observations and/or number of predictor variables. In this study, we compare logistic regression, classification tree and boosted regression models of pregnancy induced hypertension (PIH) in 19,626 patients. Predictors included age, race, parity, pre-pregnancy BMI, chronic hypertension, pregestational and gestational diabetes and smoking history. Similarities and differences in models, identification of interactions and nonlinearities, prediction accuracy, and differences in presentation of results are evaluated.

Use of Data Mining To Determine Variable Significance and Inter-relationships in Models

❖William K. Sieber, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, MS R17, Cincinnati, OH 45226, *wks1@cdc.gov*

Key Words: data mining, model, epidemiology, hearing protection

The use of data mining techniques to determine appropriate variables and interactions to investigate in a statistical or epidemiological model is demonstrated. The approach is compared with results from a model developed using more familiar techniques such as stepwise regression. Factors affecting the use of hearing protection in farmers were investigated. Age, years in farming, hearing threshold, use of noisy farm equipment, education, off-farm jobs, presence of deafness or tinnitus, training in occupational safety and health practices, and perceptions of usefulness of personal protective devices (PPD) were considered. Attitudes toward PPD, presence of tinnitus or deafness, and noisy off-farm jobs were most significant under both approaches. The data mining approach offers a graphical depiction of variables while also suggesting categorical breakdowns for continuous variables.

An Empirical Evaluation of the Random Forests Classifier Models for Variable Selection in a Large-Scale Lung Cancer Case Control Study

❖Qing Zhang, The University of Texas M.D. Anderson Cancer Center, 2815 Spring Lakes, Missouri City, TX 77459, *qzhang@mdanderson.org*; Christopher I. Amos, The University of Texas M.D. Anderson Cancer Center

Key Words: Random Forests, classification, machine learning, variable selection

Random Forests is a machine learning-based classification algorithm developed by Leo Breiman and Adele Cutler for complex data analysis. Previous research has indicated that it has excellent statistical properties when predictors are noisy and the number of variables is much larger than the number of observations. This study conducted an empirical evaluation of the method of Random Forests for variable selection using data from a large-scale lung cancer case-control study. A novel way of variable selection was proposed to automatically select prognostic factors without being adversely affected by multiple colinearities. This empirical study demonstrated that Random Forests can deal effectively and accurately with a large number of predictors simultaneously without overfitting.

S-PLUS and R Package for Least Angle Regression

❖Tim Hesterberg, Insightful Corporation, 1700 Westlake Ave N Suite 500, Seattle, WA 98109-3044, *timh@insightful.com*

Key Words: regression, regularization, L1 penalty, variable selection

Least Angle Regression is a promising new technique for variable selection applications, offering a nice alternative to stepwise regression. It provides an explanation for the similar behavior of Lasso (L1-penalized regression) and forward stagewise regression, and provides a fast implementation of both. We'll demonstrate a prototype open-source S-PLUS/R package "glars" for generalized least angle regression, extending the work by outside collaborators. See *www.insightful.com/Hesterberg/glars.*

Model and Survey Performance Measurement by the RSESP

❖James Knaub, Energy Information Administration, 2454 N Jefferson St, Arlington, VA 22207, *james.knaub@eia.doe.gov*

Key Words: regression, total survey error, performance measure, establishment survey, classical ratio estimator, cutoff sampling

The model-based relative standard error for a superpopulation, the model-based RSESP, found in Knaub (2002, 2003, and 2004), may be used to compare competing regression models for the same data sets, when regressor data are available. This tool may then be used to track survey performance to indicate changes in total survey error. Here a series of graphical results are presented and interpreted.

Mixture Model Analysis of Age-Appropriate and Delayed Vaccination

❖Andrew Baughman, Centers for Disease Control and Prevention, 1043 N Virginia Ave NE, Atlanta, GA 30306, *alb1@cdc.gov*; Charles E. Rose, Centers for Disease Control and Prevention; Kate M. Shaw, Centers for Disease Control and Prevention; Gustavo H. Dayan, Centers for Disease Control and Prevention

Key Words: mixture model, age-appropriate vaccination, Weibull distribution, health surveys

Assessment of delayed vaccination is important to identify patients at risk for disease and to plan interventions aimed at preventing outbreaks. Risk factors and interventions may differ between age-appropriate vaccination (AV) (vaccine dose was administered by the recommended age or not) and days of delayed vaccination (DV) (number of days of delay after the recommended age before the vaccine dose was administered). To simultaneously explore risk factors for both AV and DV, we developed a mixture model that contained a point distribution to model the AV binary outcome and a Weibull distribution to model time-to-vaccination for children with delayed or no vaccination. Covariates were incorporated in each distribution to quantify risk factors. The model was illustrated using data on 969 children aged 13–59 months from a vaccination coverage survey conducted in Argentina in 2002.

Applying Propensity Score Calibration To Reduce Unmeasured Confounding Bias

❖Rachel Bittner, University of Washington, 6150 NE Radford DR Apt 1112, Seattle, WA 98115, *rbittner@u.washington.edu*; Jennifer Nelson, Group Health Center for Health Studies; Michael Jackson, Group Health Center for Health Studies; Noel Weiss, University of Washington; Lisa Jackson, Group Health Center for Health Studies

Key Words: observational study, bias, public health, influenza, vaccine effectiveness, elderly

Large health care utilization databases are increasingly being used to study treatment effectiveness. But databases often contain only crude confounder measures, which may lead to residual confounding bias. Since collecting additional, more detailed confounder data is costly, it is useful to consider methods that involve collecting such data on a subset. We apply one such method, propensity score calibration (PSC) in a study of influenza vaccine efficacy (VE). The method 1) uses the full cohort to estimate a crude PS (based on crude confounders), 2) uses validation data to estimate a crude PS, a gold standard PS (based on accurate confounders), and a calibration model between them, and 3) performs regression calibration to correct the naïve VE estimate. We evaluate the ability of PSC to reduce bias using the 'control' period prior to influenza season when true VE is known to be zero.

A Case Study of Time Scale for Use in Cox's Proportional Hazard Model

❖Qimei He, Pacific Health Research Institute, 846 S Hotel Street, Suite 301, Honolulu, HI 96813, *qhe@phrihawaii.org*; Bradley Willcox, Pacific Health Research Institute; David J. Curb, Pacific Health Research Institute

Key Words: PH regression model, time scale, follow-up study, Biases

Cox's proportional hazard (PH) regression model is a widely used tool in the analysis of survival data. Most of the time, follow-up time is used as the time scale in the PH model. However, if the covariate of interest is correlated with age, the PH model using follow-up time as the time scale will produce biased results, even after adjusting for age at baseline. The Honolulu Heart Program has followed 8,006 Japanese-American in Honolulu, Hawaii for over four decades. We used the PH model to explore the relation between height and mortality in the HHP and to illustrate the time scale issue in the Cox regression model. We found that height is positively associated with mortality when using age as the time scale; conversely, a negative association is seen with mortality when follow-up time is used as the time scale, using 20-years of follow-up.

Application of the Pattern-Mixture Latent Trajectory Model in an Epidemiological Study

❖Hiroko Dodge, Oregon State University, 3066 Willakenzie Road, Eugene, OR 97401, *dodge@oregonstate.edu*; Changyu Shen, Indiana University; Mary Ganguli, University of Pittsburgh

Key Words: Latent Trajectory Model, non-ignorable missing data bias, PROC TRAJ

Longitudinal designs, requiring follow-up of the same individuals over time, are increasingly common in epidemiological studies. However, missing data bias is a major problem in longitudinal studies where attrition is inevitable over time. Restricting analyses to only the observed data could bias the results depending on the types of missingness. One approach to address non-ignorable missing data bias is a pattern mixture model, but un-identifiability is a problem. We offer a practical solution to this problem by using latent trajectory analysis implemented in the SAS TRAJ procedure, which identifies latent groups with different trajectory patterns. The approach presented here is appealing since it can be easily implemented using common software and can be applied to wide variety of disciplines which analyze longitudinal data with potentially non-ignorable missing data.

Parametric Time Dependent SIR Model

❖Di Shi, Riverdale Country School, 2700 Henry Hudson PKWY Apt 4A, Riverdale, NY 10463, *deandishi@yahoo.com*; Peter Gross, Riverdale Country School

Key Words: Parametric, Disease Spread, SIR Model

Severe Acute Respiratory Syndrome (SARS) broke out in 2003 resulting in 8098 SARS cases and 774 deaths. This epidemic caused significant social and economic disruption in the areas which it affected. While much has been learned about this syndrome since March 2003, a quantitative description of the disease spread out over time remains incomplete. A parametric time dependent disease spread SIR model was proposed in this project to study the impact of intervention. Parameters were estimated using SARS data reported by the WHO. This model can be used in simulating the effects of interventions in future epidemics.

The Association Between Anticyclic Citrullinated Peptide Antibodies and Risk of Rheumatoid Arthritis

❖Lori B. Chibnik, Brigham and Women's Hospital, 75 Frances Street PBB-B3, Boston, MA 02115, *lchibnik@rics.bwh.harvard.edu*; Lisa A. Mandl, Hospital for Special Surgery; Karen H. Costenbader, Brigham and Women's Hospital; Peter Schur, Brigham and Women's Hospital; Elizabeth W. Karlson, Brigham and Women's Hospital

Key Words: sensitivity, specificity, biomarker, Arthritis, Rheumatology

Anticyclic citrullinated peptide antibodies (CCP) are strongly associated with an increased risk of rheumatoid arthritis (RA). It is unclear if the manufacturer's recommended cut-off is optimal, or if continuous values are useful. Using two large prospective Nurses Health Study Cohorts, we examined sensitivity, specificity and mean time to RA onset at five thresholds of CCP, and also evaluated the utility of continuous titers. Incident RA was confirmed in 93 women, with three matched controls/case. Plasma was tested for CCP with the DIASTATTM ELISA assay. Results showed a threshold lower than recommended has superior test performance, higher sensitivity but low specificity. Higher titers are strongly associated with a shorter time to disease onset al.though dichotomized results are clinically meaningful, higher ranges of the titer are informative in predicting sooner development of RA.

Monte Carlo Sensitivity Analysis of Smoking as an Unmeasured Confounder in an Investigation of the Association Between Occupation and Lung Cancer Mortality

❖Jia Li, Constella Group, LLC, 5555 Ridge Ave MS R18, Cincinnati, OH 45213, *qzl0@cdc.gov*; Toni Alterman, National Institute for Occupational Safety and Health

Key Words: sensitivity analysis, Monte Carlo simulation, lung cancer mortality, smoking, occupation

Traditional statistical methods focus on measured confounders and random errors. In observational epidemiology, systematic errors due to unmeasured confounders can limit the inferential power of the study, and hence need to be addressed and quantified. Data from the National Occupational Mortality Surveillance System was used to examine the association between death from lung cancer and occupation. After adjustment for gender, race, and age, the lung cancer mortality odds ratio (MOR) and 95% confidence interval (CI) for blue-collar vs. white-collar workers was 1.25 (1.19, 1.31). Smoking status was constructed repeatedly using Monte Carlo simulation based upon associations between lung cancer mortality, smoking, and occupation obtained from the literature and the U.S. National Health Interview Survey. With additional adjustment for smoking, the MOR and 95% CI became 0.99 (0.94, 1.04).

A New Measure of Excess Longevity in Families

❖Kenneth Boucher, Huntsman Cancer Institute, 2000 Circle of Hope Dr, Salt Lake City, UT 84112, *ken.boucher@hci.utah.edu*; Richard A. Kerber, Huntsman Cancer Institute

Key Words: Familial Excess Longevity, genetics, aging

In this paper we propose a new measure of longevity in families, Modified Familial Excess Longevity (MFEL), and compare it to Familial Excess Longevity (FEL). Both measures can be readily computed using data from large genealogical databases. Familial Excess Longevity (FEL) is a weighted average of the excess longevity of the genetic relatives of a subject, weighted by the degree of relationship. FEL has been shown to correlate positively with other measures of genetic association, such as the risk of recurrence in siblings. MFEL is a related measure of familial longevity constructed using a simple regression model. The performance of MFEL is shown to be superior when both are evaluated using data from a large genealogical database containing millions of records.

A Study on the Effect of Preferred Drug Lists in Medicaid on the Overall Medical Expense

❖Joseph Richardson, Auburn University, 1547 Emily Avenue, Auburn, AL 36830, *j.richardson@mac.com*; Mark Carpenter, Auburn University; Victoria Jordan, Auburn University

Key Words: change-point regression, autoregression, Chow tests, Durbin-Watson

To gain control over the ever increasing costs associated with pharmaceuticals, especially resulting from those individuals with chronic illnesses, Medicaid Agencies often implement Preferred Drug Lists (PDL). In such programs, Medicaid patients can only select their common medications, without special permission, from this PDL. The Alabama State Medicaid Agency's current PDL program has been shown to be effective with respect to the overall pharmaceutical costs. However, the agency wanted to make sure that that the PDL does no have an adverse effect on total medical or non-pharmaceutical costs. For example, would a specified PDL lead to an increase in disease related emergency room visits?

Evaluating SEER-Medicare Linked Database Completeness for Oral and Pharyngeal Cancer

Jonathan Mahnken, The University of Kansas Medical Center; ❖John Keighley, The University of Kansas Medical Center, 3901 Rainbow Blvd, MSN 1008, Kansas City, KS 66160, *jkeighle@kumc.edu*; Matthew Mayo, The University of Kansas Medical Center

Key Words: secondary data analysis, health outcomes research, Poisson Regression

Oral and pharyngeal (OP) cancer is a persistent oral health problem. Baseline and trend data to measure progress are lacking. Our long-term goal is to create an algorithm using Medicare claims to identify OP cancer cases among those 65+. The goal of this project was to assess the completeness of the SEER-Medicare linked database (SEER-Med). We compared IRs from the "gold-standard" SEER public use database (SEER) to those from SEER-Med using Poisson regression, adjusted for age, sex, race, and ethnicity (and interactions with data source). The IR (95% CI) for OP cancer (/100,000 person-years) using SEER-Med was 31.2 (30.6-31.8), and from SEER was 43.0 (42.3-43.7). Thus, about 1/4 of the cases were missing from SEER-Med (IRR: 0.73; 0.71-0.74). Significant interactions indicated completeness varied by race (p< 0.0001) and ethnicity (p< 0.0001).

Cost-Effectiveness Analysis in Breast Cancer Early Detection

❖Charlotte H. Ahern, Rice University, 6100 Main Street, MS-138, Houston, TX 77005, *hsiehc@rice.edu*; Yu Shen, The University of Texas M.D. Anderson Cancer Center

Key Words: early detection, breast cancer, cancer screening, cost-effectiveness

Breast cancer screening strategies combining mammography and clinical breast exam are considered to be effective in the early detection of breast cancer. In this study, we assess the cost-effectiveness of various screening strategies combining these two modalities. We account for factors that impact the performance of a screening program including age-specific incidence, pre-clinical duration of the disease, screening sensitivity and specificity, and treatments. The outcomes of interest are quality-adjusted years of life saved and incurred costs. We compare ten strategies in a comprehensive Monte Carlo simulation, taking into consideration costs involving screening exams, biopsy caused by false-positive exams and subsequent treatments. Our results indicate that a cost-effective screening program gives mammography and clinical breast exam in alternating years from ages 40 to 79.

What Do We Know About the Effects of Substance Use on Labor Productivity? A Survey of Results and Related Methodological Issues

❖Richard Bryant, University of Missouri-Rolla, Economics and Finance Department, Rolla, MO 65401, *rrbryant@umr.edu*; V. A. Samaranayake, University of Missouri-Rolla

Key Words: Alcohol and Drug Abuse, Wages, Hours Worked, Causal Inference, Longitudinal Models

The effect of substance abuse on productivity is an important issue in public policy. This topic has received considerable attention since the 1980s, with numerous publications reporting estimates of the economic impact of substance abuse. These results, however, do not paint a consistent picture. Some studies imply negative impacts of substance abuse on productivity, while others show a positive relationship. It is not clear that such discrepancies can be explained by differences in the data. The divergent results can be due to methodological differences. Further, the statistical associations found by researchers may not be indicative of a causal relationship. While many authors have resorted to longitudinal models to account for unobserved factors, it is not clear that this approach leads to true causal inference. We investigate this issue and explore possible solutions.

School Social Support and Adolescent Depressiveness: Multilevel Analysis of Finnish School Health Promotion Survey

❖ Noora Ellonen, University of Tampere, Harmalansaarenkatu 6 A 11, Tampere, FIN-33900 Finland, *noora.ellonen@uta.fi*

Key Words: adolescent, depressiveness, school effect, multilevel analysis, social support, climate consensus

This study elaborates the meaning of the school social support in adolescent mental health. Adolescent depressiveness is explained with school social support provided by teachers and classmates and operationalized as quality (average) and consensus (variation) of the climate. The analysis is made with multilevel logistic regression analysis, which is conducted with MLwiN 1.1 and the parameters are estimated with the RIGLS method. Results of the analysis suggest that the amount of school social support have no significant effect on adolescent depressiveness, although at the individual-level the social support perception shows a meaningful effect. The consensus about the social climate, however, has a significant increasing effect on adolescent depressiveness. Analysis is based on School Promotion Health Survey, which covers about 90 % of Finnish 14-16-years-old adolescents.

Bayesian Multiple Imputation and Maximum Likelihood Methods for Missing Data

❖ Min Sun, Sam Houston State University, 1701 Bobby K Marks Dr, Apt C100, Huntsville, TX 77340, *stdmxs19@shsu.edu*; Ferry Butar Butar, Sam Houston State University

Key Words: multiple imputation, Bayesian method, maximum likelihood, monte carlo, boostrap

Bayesian multiple imputation (MI) and Maximum Likelihood (ML) provides a useful strategy for dealing with datasets included missing values. Imputation methods affect the significance of test results and the quality of estimates. In this paper, the general procedures of MI and ML described, which include the normal-based analysis of a multiple imputed dataset. A Monte Carlo simulation is conducted to compare the performances of the methods.

Imputation for Missing Physiological Measurement Data: Simulations and Applications

❖ Matt Jans, University of Michigan, 410 Nob Hill Ct Apt 2, Ann Arbor, MI 48103, *mattjans@isr.umich.edu*; Steven Heeringa, Institute for Social Research

Key Words: Biological measurements, missing data, multiple imputation

This poster will address alternative approaches to item missing data imputation in univariate and multivariate regression analysis involving physiological measurements collected in health surveys. The National Health and Nutrition Examination Survey (NHANES II) data file includes data on 20,322 respondents from a nationally representative stratified cluster sample. The NHANES II data on blood pressure measurements will be used as the basis for a simulation study that will examine the properties of several imputation methods (including single and multiple imputation) to address missing data problems generated under plausible missing at random (MAR) and non-ignorable mechanisms for physiological measurement data. Comparative results will also be presented for an empirical investigation of imputation outcomes for missing blood chemistry measures from the 2002 Chicago Mind and Body Study.

A Comparison of Weighting Adjustment and Multiple Imputation Methods To Correct for Nonresponse Bias in a Longitudinal Group-Randomized Clinical Trial for Depression

❖ Lingqi Tang, University of California, Los Angeles, 10920 Wilshire Blvd Suite 300, Health Services Research Center, Los Angeles, CA 90024-6505, *lqtang@ucla.edu*; Naihua Duan, University of California, Los Angeles; Ruth Klap, University of California, Los Angeles; Thomas Belin, University of California, Los Angeles

Key Words: Missing Data, Hot Deck, Nornesponse weights, Multiple Imputation, Model-Based Imputation, Logistic regression

Weighting adjustment is a standard method used to correct for unit nonresponse bias. However, in long term follow-up, the association between auxiliary variables and the outcome variables may not be strong. In this paper, we compare two methods of handling unit nonresponse between a baseline assessment and 9-year follow-up in a group-level randomized controlled trial of quality improvement for depression treatment. One approach uses nonresponse weights constructed by fitting logistic regression models to predict follow-up status from baseline clinical and sociodemographic characteristics. The reciprocal of the predicted follow-up probability is used as the nonresponse weight for each participant. A second method uses multiple imputation based on the approximate Bayesian bootstrap. Results from the two approaches are compared.

Comparison of Modeling Approaches to Deriving Normative Values of Bone Mineral Density from NHANES III

❖ Qing Wu, Tulane University, *qingwu2000@yahoo.com*; John Lefante, Tulane University; Jeanette Magnus, Tulane University; Janet Rice, Tulane University

Key Words: Bone Mineral Density, NHANES III, osteoporosis, Model

Developing accurate normative reference bone mineral density (BMD) values is important to diagnose osteoporosis and prevent fractures. The study aim was to compare linear and polynomial regression models by residual analysis to identify the best model to derive normative BMD values at various skeletal sites using NHANES III. Gender- and race-specific group data analyses were conducted on 9779 aged 20–65 subjects who had no conditions or medication affecting bone metabolisms. Sampling and design effects were addressed using STATA 9.1. The results show that polynomial regression equations provide a significantly better fit than linear equations. These race- and gender- specific polynomial regression models, with their respective lower limits of normal, have the potential to improve diagnostic accuracy for low bone mass and can be use for populations other than postmenopausal women.

Assessing the Potential Impact of Missing Data in a Longitudinal Study with a Continuous Outcome

❖Owen Devine, Centers for Disease Control and Prevention, 1600 Clifton Road, MS E-87, Atlanta, GA 30333, *ojd1@cdc.gov*; Jorge Rosenthall, Centers for Disease Control and Prevention; Nyasha Skerrette, Emory University

Key Words: Missing Data, Uncertainty, Sensitivity

Based on Daniels and Hogan (Biometrics, 2000), we employ a simple method to assess the sensitivity of model-based parameter estimates developed using study completers to assumptions on possible outcomes for dropouts. A key issue in this type of assessment is generating reasonable potential values for missing data given available information. We develop uncertainty distributions, based on expert opinion and/or analysis of available data, for parameters underlying the missing data. Monte Carlo simulation is used to generate realizations for both the missing data and the associated model parameter estimates that reflect both uncertainty in the underlying parameters and sampling variation inherent in model estimation. We illustrate the approach using a longitudinal study comparing the effectiveness of two folic acid dosing schemes among Honduran women.

Area Under the Curve and Other Summary Indicators of Repeated Waking Cortisol Measurements

❖Desta Fekedulegn, National Institute for Occupational Safety and Health, Health Effects Laboratory Division, Morgantown, WV 26505, *djf7@cdc.gov*; Michael Andrew, Centers for Disease Control and Prevention; Cecil M. Burchfiel, National Institute for Occupational Safety and Health; John M. Violanti, State University of New York at Buffalo; Tara A. Hartley, National Institute for Occupational Safety and Health; Luenda E. Charles, National Institute for Occupational Safety and Health; Diane B. Miller, National Institute for Occupational Safety and Health

Key Words: Repeated cortisol measurements, Area under the curve, principal component analysis, Total hormonal secretion, Time course of salivary cortisol

In research designed to assess health consequences of stress, repeated measurements of salivary cortisol are often used as a physiologic indicator of the responsiveness of the hypothalamic-pituitary-adrenal axis. There is a need to derive two forms of area under the curve that summarize the repeated measurements: AUCG and AUCI. In this study interpretation and generic definition of the area under the curve was provided through graphical analyses and examination of its association with other summary measures. In generic form AUCI is the area under the curve above the baseline value minus the area above the curve below the baseline value. The sign and magnitude of the parameter are related to the profile and the rate of change of the measurements over time. Principal components analyses revealed two groups of summary measures: measures of magnitude of response and pattern of response.

On the Meta-Analysis of Incomplete Primary Study Data

❖Todd Bodner, Portland State University, 5744 SE Preston Ct, Hillsboro, OR 97123, *tbodner@pdx.edu*

Key Words: Meta-analysis, Missing data, Effect sizes

Meta-analysis serves to compare and combine primary study results often stated as effect sizes. However, incomplete data are common in primary research studies. Despite large literatures on meta-analytic and missing data methods, little research has explored the impact of incomplete primary study data on effect sizes and their weights and subsequently on meta-analytic inference under the variety of missing data methods. This poster focuses on the computation of effect sizes and their weights along with their implications for meta-analytic inference under situations with complete and varying degrees of incomplete data handled using either multiple imputation or listwise deletion. Results include the finding that under MCAR these missing data methods can generate large differences in effect size weights depending on the amount of missing data which can impact meta-analytic inference.

Response and Nonresponse Pattern Analysis in Survey Research

❖Chong Ho Yu, Arizona State University, 3S89 Computing Commons, Tempe, AZ 85287, *chonghoyu@gmail.com*; Samuel DiGangi, Arizona State University; Sandra Andrews, Arizona State University; Angel Jannasch-Pennell, Arizona State University

Key Words: survey, data mining, classification tree

In a southwestern university, a survey regarding use of instructional technology was sent to all students via email. Respondents in survey research were self-selected and thus parametric procedures might not be properly applied. While it is important to analyze the data collected from the respondents, it is equally important to examine the profile of non-respondents in order to understand the potential bias of the survey results. To address this issue, the demographic information of the entire population was extracted from the data warehouse for response pattern analysis using sata visualization and data mining techniques. It was found that contrary to prior research, gender and race were not significant predictors to survey responses. Rather, science and engineering students tended to respond to this type of technology-related survey.

Standard Error Estimation in Latent Variable Model

❖Yan Wang, Southern Methodist University, Dept. of Statistical Science, 144 Heroy Science Hall, 3225 Daniel Avenue, Dallas, TX 75275-0332, *yanw@smu.edu*; S. Lynne Stokes, Southern Methodist University; Andreas Oranje, Educational Testing Service

Key Words: Latent Variable Model, Standard Error, NAEP

In the large-scale educational assessment, the observations are selected from population by the complex sampling design, and the quantity of interest, such as student ability, is not observed directly. The data selection method causes clustering in the observations which are often positively correlated. The manifest variables of the student ability in the educational measurement are commonly the student's responses to a group of questions. The questions are often related to the same stimulus which causes positive correlations among the student's responses. Ignoring the positive correlation and making independence assumption in the standard statistical analysis underestimate the standard errors of the estimates. The new models relax two sources of the independences mentioned above and are used to obtain the proper standard errors. We investigate the standard error by simulation studies.

Nonlinear Random Coefficients Regression for Estimating an Interaction Term

Reid Landes, University of Arkansas for Medical Sciences; ❖Jeffrey A. Pitcock, University of Arkansas for Medical Sciences, 4301 W Markham St, Slot 850, Little Rock, AR 72205, *pitcockjefferya@uams.edu*

Key Words: delay discounting, impulsivity, clinical trial, heterogeneity of variance

Many clinical trials are designed to compare how much a standard and experimental population change between pre- and post-treatment. In other words, the interaction of treatment and time—a difference of differences—is of interest. We consider how to make this comparison when an individual's outcome of interest is estimated with nonlinear regression. In particular, we describe how to use all of the raw data in a nonlinear random coefficients regression that allows direct estimation of the difference of differences. Further, we also illustrate how to account for heterogeneity of

variance among the individuals. We apply this method to impulsivity data collected from treatment-seeking opiate addicts in a two-treatment clinical trial. The results from the full nonlinear regression are compared with ad hoc methods which are typically used for these sorts of problems.

Relationships Between Students' Science Achievement, Attitudes Toward Science, and Learning Through Inquiry

❖ Man Hung, University of Utah, 161 University Village, Salt Lake City, UT 84108, *mh1238@hotmail.com*

Key Words: science education, SEM, inquiry, attitudes, achievement, TIMSS

With the overarching goal of improving scientific literacy for all, the emphasis on science achievement has been expanding all over the world. In this study, I seek to identify the relationships among learning through inquiry, attitudes toward science and students' science performance in the United States. Data for this study came from the Trend in International Mathematics and Science Study (TIMSS) 2003. Analyses conducted here mainly focused on the 8th graders' composite science scores in the achievement test and their relationships with the selected variables in the student questionnaire. There were a total of 8,912 8th graders participated in TIMSS 2003 in the United States. Based on prior research, a series of structural equation models were constructed and tested and their results were analyzed. Implications for future research were also discussed.

Evaluating a Multilevel Item Response Theory Model

❖ Thomas Schmitt, University of Wisconsin-Milwaukee, 3215A S New York Ave, Milwaukee, WI 53207, *schmitta@uwm.edu*; Cindy M. Walker, University of Wisconsin-Milwaukee

Key Words: item response theory, multilevel models

Embedding item response theory (IRT) models within multilevel models tends to improve parameter estimates over an approach in which the latent trait is first estimated and then used as an outcome variable in a common linear model. Though extensions have been proposed, there exists a dearth of research systematically investigating multilevel IRT model's behavior under different conditions and when the model fails to provide robust parameter estimates. A simulation study will be conducted to evaluate parameter estimation error of a three-level IRT model under the following conditions: sample size at level-2 and level-3, the number of predictors at level-2, the number of items at level-1, unequal variances at level-3, and the shape of the latent trait distribution. The results will be used to provide researchers with a set of recommendations regarding the investigated conditions.

Globalization and Chinese City System Restructuring

❖ Xiulian Ma, The University of Utah, 1156 Heritage Center, Salt Lake City, UT 84112, *xiulian.ma@soc.utah.edu*

The paper will apply a random graph model in social network analysis to examine to what extent global forces have restructured Chinese city system. The random graph model can solve the problem that in network data observations are not independent to each other. A country's city system often entails certain territorial integration function, and city size distribution (like rank-size rule or primate) can facilitate or arrest the country's economic development. According to various theoretical perspectives, under globalization, Chinese city system will become more hierarchical, especially with the rise of primate cities. To what extent is this argument true? Has globalization significantly restructured Chinese city system? To answer these questions, the study will examine attributes of the 40 largest cities in China (1992–2003).

210 ASA President's Invited Address

The ASA, IMS, ENAR, WNAR, SSC

Monday, July 30, 4:00 pm–5:50 pm

Regularization Methods in Statistical Model-Building: Statisticians, Computer Scientists, Classification, and Machine Learning

❖ Grace Wahba, University of Wisconsin-Madison, Statistics Department, Madison, WI 53705, *wahba@stat.wisc.edu*

Key Words: regularization, cross validation, model building, optimization

We survey members of a broad class of statistical model building tools that are popular in nonparametric regression and classification, which have the common the feature that they involve an optimization problem with an explicit tradeoff between fit to the data and complexity of the model. We examine some relationships between Bayes estimates, penalized likelihood nonparametric regression methods, and the classification method known as a Support Vector Machine, in the context they share as regularization methods. Cross validation based methods for choosing the tradeoff (tuning) parameters will be examined, along with problems in selecting important variables and variable clusters. Interplay between Statisticians and Computer Scientists in extending this rich class of methods will be shown to be valuable to both.

211 IMS Presidential Address

IMS

Monday, July 30, 8:00 pm–9:30 pm

Open Access to Professional Information

❖ Jim Pitman, University of California, Berkeley, Department of Statistics, Berkeley, CA 94720-3860, *pitman@stat.berkeley.edu*

A number of online resources now provide open access to information of value to students and professionals in probability and statistics. These include arXiv, Google Scholar, Wikipedia, MathWorld, PlanetMath, and several electronic journals. These developments are perceived by some professional societies as threats to their business model. By contrast, IMS sees open access electronic resources as effective means of achieving its primary purpose: to foster the development and dissemination of the theory and applications of statistics and probability.

212 ASA Council of Chapters Roundtable with Coffee (fee event)

Council of Chapters

Tuesday, July 31, 7:00 am–8:15 am

AP Statistics and Local Chapters of the ASA

❖ Dongseok Choi, Oregon Health & Science University, 3181 SW Sam Jackson Park Road CB669, Portland, OR 97201, *choid@ohsu.edu*

Key Words: AP-statistics, ASA Chapters, Partnership

The Advanced Placement Program (AP) offers courses including statistics. Each AP course has a corresponding exam administered in May. Most colleges and universities in the US have an AP policy granting incoming students credit, placement, or both on the basis of their AP exam grades. The AP statistics standard curriculum is equivalent to a noncalculus, one-semester introductory college course in statistics. The Oregon Chapter of The ASA partnered with a few teachers for AP statistics, and some common challenges and needs were identified, e.g., shortage of properly trained teachers, lack of real-world examples, wishing for opportunities for interactions between teachers and professional statisticians, and class visits by statisticians. In this session, we will discuss how ASA Chapters can be resources for teachers and students in AP statistics and share our experiences.

213 Section on Health Policy Statistics Roundtable with Coffee (fee event)

Section on Health Policy Statistics

Tuesday, July 31, 7:00 am–8:15 am

Approaches to Measuring Differences in Health That Are Unaffected by the Prevalence of an Outcome

❖James P. Scanlan, Attorney at Law, 1529 Wisconsin Ave NW, Suite 300, Washington, DC 20007, *jps@jpscanlan.com*

Key Words: health disparities, binary variables, continuous variables, mortality differences, health inequality, measurement

Binary measures of group differences in health outcomes vary, depending on the prevalence of the outcome. For example, as mortality declines, relative differences in mortality tend to increase, while relative differences in survival tend to decline. Measures that are functions of binary outcomes (e.g., gini coefficients, longevity differences) also tend to change in some manner as the prevalence of an outcome changes. This renders problematic efforts to evaluate the size of health disparities at different times or in different settings with measures based on mortality or morbidity rates. It is possible that continuous variables may offer approaches to measuring the size of group differences in health that are independent of changes in prevalence of the factor being examined. But that possibility must be tested, and factors that are useful to examine need to be identified.

214 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education

Tuesday, July 31, 7:00 am–8:15 am

Coordinating the Large Undergraduate Course

❖Roger Woodard, North Carolina State University, Box 8203 2501 Founders Drive, Raleigh, NC 27695, *woodard@stat.ncsu.edu*

Key Words: coordination, TA, exams, large classes

This roundtable will address coordination of large undergraduate courses. Discussion will concentrate on practical solutions to administrative problems and improving student learning. Issues addressed will include working with large numbers of TAs, administering exams to large numbers of students, and handling general administrative tasks.

215 Section on Statistical Graphics Roundtable with Coffee (fee event)

Section on Statistical Graphics

Tuesday, July 31, 7:00 am–8:15 am

Network Visualization

❖Deborah Swayne, AT&T Labs - Research, Room B295 Shannon Laboratory, 180 Park Avenue, Florham Park, NJ 07932-0971, *dfswayne@yahoo.com*

Key Words: graph, network, visualization, interaction, graphics

The group will discuss the exploration of network data, by which we mean data that includes a mathematical graph, a set of vertices connected by edges. A key starting point is graph layout. We will discuss attributes of a good graph layout, useful web sites, and publicly available layout software. We also will discuss interaction with such graphs and will look at a variety of interesting pictures.

216 Section on Statistics and Marketing Roundtable with Coffee (fee event)

Section on Statistics and Marketing

Tuesday, July 31, 7:00 am–8:15 am

Statistics in Humor and Art

❖Igor Mandel, *igor.mandel@yahoo.com*

Key Words: Art, Humor, Marketing Statistics

At first blush, statistics do not seem to have much in common with humor and the arts. However, there are several examples from art, poetry, literature, and film where statistics are used as a theme. Many of us in our everyday work lives use humor to impart difficult statistical concepts to our clients in ways they can understand. Join us in discussing how humor, the arts and statistics intersect. Bring your favorite statistical pictures, poems, web sites, etc.

217 Section on Survey Research Methods Roundtable with Coffee (fee event)

Section on Survey Research Methods

Tuesday, July 31, 7:00 am–8:15 am

The Transition from Graduate School to a Career in Statistics

❖Edward English, NORC at the University of Chicago, 55 E Monroe St, Suite 4800, Chicago, IL 60603, *english-ned@norc.org*

Key Words: Career, Job, Entry level, Survey research

The purpose of this roundtable is to discuss common issues and situations that arise for statisticians and related professionals when transitioning from graduate school to the “real world.” Topics to be considered include: networking, job interviews, financial realities, workplace dynamics, and career development. An underlying theme concerns how to understand the similarities and differences between academic and work environments to overcome common challenges. The moderator in this case has six years of professional experience after graduate school, so can offer his perspective. It is intended that participants be able to share their own experiences with the group in a conversational setting.

218 Section on Teaching Statistics in the Health Sciences Roundtables with Coffee (fee event)

Section on Teaching Statistics in the Health Sciences
Tuesday, July 31, 7:00 am–8:15 am

Ideas for Improving Appreciation and Understanding of Cluster Sampling

❖Winston A. Richards, The Pennsylvania State University, Harrisburg, 2100 Chestnut Street, Harrisburg, PA 17104, *ugu@psu.edu*

Key Words: Simple Random Sampling, Cluster Sampling, Confidence intervals

Health researchers must master basic principles and methods for critiquing and producing health research literature efficiently; thus, they need to understand various methods of sampling. Two recent articles in the February 2007 issue of The American Statistician, one by Alf and Lohr, the other by Strasak et al, can motivate discussion about how to introduce these concepts to students in the health sciences. Al and Lohr review the treatment of sampling assumptions for confidence intervals and hypothesis tests in several texts. They identify strengths and flaws and propose promising approaches to teaching simple random sampling, cluster sampling and confidence intervals. Strasak et al. identify related errors that show up in medical research articles. They provide a quality improvement checklist for statistical evaluation of medical manuscripts.

Interpreting Statistical Results from Medical Literature in the Classroom

❖Kirk Anderson, Grand Valley State University, 1 Campus Drive, Allendale, MI 49401, *anderkir@gvsu.edu*

Key Words: medical literature, health sciences, statistics education

If you are considering the use of medical literature in your classes, or have experiences to share, please join us for coffee and discussion. Issues may include how/where to look for articles, who should conduct the search, how to handle mistakes or complicated methods, and what to ask to assess student comprehension. I am a statistics professor with experience teaching introductory statistics to graduate students who are typically biologists, nurses, occupational therapists, physical therapists, and physician assistants.

219 Introductory Overview Lecture 3: Improving Statistical Presentations and Posters

The ASA, SSC, ENAR, IMS, WNAR
Tuesday, July 31, 8:30 am–10:20 am

Improving Statistical Presentations

❖Mike Colopy, GlaxoSmithKline, 223 Larkwood Lane, Cary, NC 27518, *mike.w.colopy@gsk.com*

Good presentations skills enhance one’s ability to communicate ideas and enhance professional growth. The objective of this presentation is to share tried and true techniques for (1) avoiding common pitfalls of technical presentations, (2) developing a captivating presentation, and (3) effectively conveying verbal and nonverbal messages.

Improving Statistical Posters

❖Dianne Cook, Iowa State University, Dept of Statistics, 325 Snedecor Hall, Ames, IA 50011-1210, *dicook@iastate.edu*

Poster presentations provide an opportunity to communicate at a more intimate, flexible level than oral presentations. This presentation will discuss ways to structure a printed poster, and use technology, to invite readers into your work and provoke questions. Some of the thoughts are based on our own experience, with the JSM 2006 data expo competition, and some material will be drawn from the literature on visual communication.

220 Practical Designs in Early-Phase Cancer Trials ●

ENAR, Biopharmaceutical Section, WNAR
Tuesday, July 31, 8:30 am–10:20 am

Comparing Two-Stage Treatment Strategies Based on Sequential Failure Times Subject to Interval Censoring

❖Peter F. Thall, The University of Texas M.D. Anderson Cancer Center, Dept. of Biostatistics, Box 447, 1515 Holcombe Blvd., Houston, TX 77030, *rex@mdanderson.org*; Leiko H. Wooten, The University of Texas M.D. Anderson Cancer Center; Nizar M. Tannir, The University of Texas M.D. Anderson Cancer Center; Randall E. Millikan, The University of Texas M.D. Anderson Cancer Center; Christopher J. Logothetis, The University of Texas M.D. Anderson Cancer Center

Key Words: Clinical trials, Dynamic treatment regime, Interval censoring, Simulation, Treatment Policy, Multi-stage therapy

For many diseases, therapy involves multiple stages, with treatment in each stage chosen based on the patient’s current disease status and history of treatments and clinical outcomes. We present a statistical framework for clinical trials comparing two-stage strategies based on time to second disease worsening or discontinuation of therapy. Each patient is randomized among a set of treatments at enrollment. If disease worsening occurs the patient is re-randomized among a set of treatments excluding the treatment given initially. The goal is to select the best two-stage strategy in terms of overall failure time. We present a model accounting for non-constant failure time hazards, regression of second failure time on first worsening time, interval censoring and delay between the two stages of therapy. The method is illustrated by a metastatic renal cancer trial.

Two-Stage Designs for Dose-Finding Trials with a Biologic Endpoint Using Stepwise Tests

❖Mei-Yin C. Polley, Amgen Inc., 2 Bayside Village Place, Suite # 301, Clinical Development Biostatistics, San Francisco, CA 94107, *meic@amgen.com*; Ying-Kuen K. Cheung, Columbia University

Key Words: Familywise error rate, Futility interim, Minimum effective dose, Monotonicity, Multiple comparison, Probability of correct selection

We tackle the problem of early phase dose-finding trials with monotone biologic endpoints such as biologic measurements and laboratory values. A specific aim of this type of trial is to identify the minimum dose that exhibits adequate drug activity and shifts the mean of the endpoint from a zero doe, the so-called minimum effective dose. Stepwise tests for dose-finding have been well studied in the context of non-human studies where the sampling plan is done in one stage. We extend the notion of stepwise tests to a two-stage setting in an attempt to reduce the sample size requirement by shutting down unpromising doses in a futility interim. Specifically, we examine four two-stage designs and apply them to design a statin trial with four doses and a placebo in patients with Hodgkin's disease. We discuss the calibration of the design parameters and the implementation of these methods.

A Strategic Approach to Designing Early Phase Cancer Trials

❖Tze Leung Lai, Stanford University, Department of Statistics, Sequoia Hall, Stanford, CA 94305-4065, *lait@stat.stanford.edu*; Philip Lavori, Stanford University; Mei-Chiung Shih, Stanford University

Key Words: adaptive design, cancer trials, safety, efficacy, futility stopping, internal pilot

Early phase cancer trials are typically small in size and relatively short in duration. There are also ethical constraints in not subjecting patients to excessive toxicity. The goals and scope of such trials are necessarily limited. We propose a strategic approach that incorporates the longer-term goal towards establishing the efficacy and safety of the new therapy in the design of early phase trials. In particular, by making use of recent advances in adaptive design and inference, our strategic approach can enable the results of the Phase II trial to be used as an internal pilot of a later Phase III trial. Early stopping for futility will also save resources if the longer-term goal is found unlikely to be met. These ideas for Phase II/III trials can also be modified for Phase I/II studies. Institutional issues in implementing these trials are also discussed.

221 Valid and Reliable Disability Measurement in Surveys ● ✪

Committee on Statistics and Disability, Social Statistics Section, Section on Survey Research Methods

Tuesday, July 31, 8:30 am–10:20 am

Identifying the Population with Disabilities: A Comparison of Current Survey Estimates

❖Andrew J. Houtenville, Cornell University, 303B ILR Extension Building, Ithaca, NY 14853, *ajh29@cornell.edu*; S. Antonio Ruiz-Quintanilla, Cornell University; Robert Weathers, Mathematica Policy Research, Inc.

Key Words: Disability, Employment, Poverty, Surveys, Population, Trends

In the United States, there is no single, universally accepted definition of disability in either the government programs that serve people with disabilities or among federal surveys. This presentation provides an overview of the national data sources that allow for the identification of the population with disabilities: ACS, NHIS, SIPP, CPS-ASEC, Census 2000, and other data sources. The strengths and weakness of each data source will be discussed, with a particular focus on the survey items used to identify the population with disabilities. Estimates derived from these data sources will be compared, including estimate of the size of the population with disabilities, the disability prevalence rate, and employment and poverty rate for people with and without disabilities. In addition, current efforts to design new sets of disability survey items and the SIPP/DEWS will be discussed.

The Development and Testing of Disability Questions for the Current Population Survey

❖Terence M. McMenamin, Bureau of Labor Statistics, 2 Massachusetts Ave. N.E., Washington, DC 20212, *mcmenamin_t@bls.gov*; Stephen M. Miller, Bureau of Labor Statistics; Anne E. Polivka, Bureau of Labor Statistics

Key Words: disability, CPS, BLS, health, DOL, survey

In accordance with Executive Order 13078, the Bureau of Labor Statistics, in cooperation with the Employment Rate Measurement Methodology interagency workgroup, identified the goal of placing a small set of questions within the Current Population Survey (CPS). A set of potential questions was drawn from existing surveys, cognitively tested, and placed in the National Comorbidity Survey (NCS) for a field test. Based on an analysis of the NCS data, a set of seven questions was identified, and cognitively tested to ensure the questions would work within the CPS. This question set was placed in the February 2006 CPS for testing. Analysis of the test data revealed a lower overall disability rate as measured in the CPS than in the NCS, with lower positive response rates for each question. The data did not indicate that adding the questions caused an adverse effect on the response rates.

Disability Data and the American Community Survey

❖Sharon Stern, U.S. Census Bureau, Washington, DC 20233, *sharon.m.stern@census.gov*

Key Words: disability statistics, survey methods, data collection, American Community Survey

The Census Bureau collected disability data in each of the last four decennial censuses. As a continuous survey collecting demographic, social, and economic characteristics at the local level, the American Community Survey will eliminate the need for a long form sample in the 2010 census. While the original ACS disability questions were designed to be similar to the Census 2000 questionnaire, evaluations of those data and re-examination of the agency needs for disability data resulted in the inclusion of disability as a subject for testing alternatives for future ACS questionnaires. This presentation discusses the evolution of the questionnaire items in response to changing agency needs and provides a review of the content test results with regard to the measurement of disability.

222 Present-Day Challenges in the Design and Analysis of Industrial Experiments ● ✪

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Tuesday, July 31, 8:30 am–10:20 am

A Structural Approach to the Design of Experiments for Robust Parameter Design

❖Raymond H. Myers, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060, *rmyers@vt.edu*

Key Words: robust design, design of experiments, statistical process control, noise factors, response surface methods

In this talk, we discuss new criteria for the design of experiments in the robust design setting.

A Posterior Predictive Approach to Multiple Response Surface Optimization

John Peterson, GlaxoSmithKline; ❖Enrique del Castillo, The Pennsylvania State University, 357 Leonhard Building, University Park, PA 16802, *exd13@psu.edu*

Key Words: batch effects, Bayesian Model Averaging, Markov Chain Monte Carlo, missing values, noise variables, robust parameter design

This presentation provides an overview of an approach to multiple response surface optimization that provides optimal operating conditions and also measures the reliability of an acceptable quality result. The traditional optimization approaches of overlapping surfaces or the desirability function do not take into account the covariance structure of the data nor the model parameter uncertainty. The proposed posterior predictive approach can be used with most of the current multiresponse optimization procedures to assess the reliability of a future response. This posterior predictive approach is easy to interpret and it takes into account the correlation structure of the data, the variability of the process distribution, and the model parameter uncertainty. This approach can also be extended to accommodate noise variables, batch effects, missing values, or model form uncertainty.

Componentwise Variance Dispersion Graphs for Mixture Experiments

❖John J. Borkowski, Montana State University, Department of Mathematical Sciences, Bozeman, MT 59715, *jobo@math.montana.edu*; Boonorm Chomtee, Kasetsart University; Kamolchanok Choochaow, Silpakorn University

Key Words: mixture experiment, variance dispersion graph, prediction variance trace, G-efficiency

Consider a q-component mixture experiment with constraints defined by lower and upper bounds for each component proportion. Whether the design region is a simplex or some irregularly shaped subregion, the problem of interest is the extension of variance dispersion graphs (VDGs) to experiments involving mixtures. As an alternative to studying the prediction variance on shrunken polyhedral spaces (Piepel and Anderson 1993) or to the prediction variance trace (PVT) plots involving Cox directions (Vining, Cornell, and Myers 1993), a new graphical tool, called the componentwise variance dispersion graph (CVDG), will be proposed. A CVDG contains componentwise plots of the variability throughout the entire polyhedral design space. A CVDG also provides the mixture yielding the maximum prediction variance which in turn provides the design's G-efficiency. Examples of CVDGs will be presented.

223 Toxicogenomics: From Concept to Regulatory Issues ● ✪

Biopharmaceutical Section, Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

Toxicogenomics for Toxicity Evaluation and Biomarker Identification

❖Hisham Hamadeh, Amgen Inc., One Amgen Center Dr, Mailstop: 14-1-B, Thousand Oaks, CA 91320, *hhamadeh@amgen.com*

Key Words: toxicity, toxicogenomics, biomarker, microarrays

The use of genomic resources to aid in toxicity evaluation is gaining more acceptance and interest within academic, governmental, and industrial settings. Attributing transcript, protein, or metabolite fluctuations as associations with, or precursors to, toxicity outcomes has been challenging. Various strategies will be presented aimed at delineating specific and sensitive toxicity-related fluctuations in transcripts, proteins, or metabolites. The discussion will extend to the identification of potential biomarkers that may be used for predictive or diagnostic purposes.

Systematic Mining of a Large Gene Expression Database and Design of a Multi-Endpoint Diagnostic Device

❖Georges Natsoulis, Datahooks, Inc., 256 Stanford Avenue, Kensington, CA 94708, *gnatsoulis@datahooks.com*

Key Words: data mining, gene expression, diagnostic device, classification, toxicity, pharmacology

A contextual database was created to relate the physiological outcomes induced by diverse compounds to the gene expression patterns measured in the same animals. This database developed from in vivo treated rats provides the context and supporting data to accelerate accurate interpretation of mechanisms of toxicity and pharmacology of chemicals and drugs. To date, approximately 600 different compounds, including more than 400 FDA approved drugs, have been profiled in up to 7 different tissues of rats. We describe the systematic mining of the DrugMatrix database using a supervised classification approach and determine the set of all resolvable endpoints. We show that a relatively small number of genes are used in a large number of signatures and that a small set of genes is sufficient to rederive signatures for all the above end-points, suggesting the basis of a diagnostic device.

Gene Signatures for Hepatotoxicants

❖Nandini Raghavan , Johnson & Johnson PRD, G-031, 1000 Rt. 202-S., Raritan, NJ , *nraghava@prdus.jnj.com*; Dhammika Amaratunga, Johnson & Johnson PRD; Alex Nie, Johnson & Johnson PRD; Michael McMillian, Johnson & Johnson PRD

Key Words: genomics, microarray, gene signatures, classification, toxicity, pharmaceutical

Toxicogenomics is an evolving area that attempts to exploit new high-throughput "omics" technologies to develop faster, better and earlier safety screens for drugs. Improving the screening process for new compounds to predict toxicity issues early in discovery will enable pharmaceutical companies to reduce the costs of developing new drugs significantly. In this talk we discuss some of the issues arising in the design and analysis of toxicogenomic experiments, as well as methods for classifying compounds into various hepatotoxicity classes. We also discuss the incorporation of genomic databases to construct biologically relevant gene signatures for predicting toxicity in new compounds, and cross-platform comparisons of gene signatures. We will also present the results of a large-scale study at Johnson and Johnson.

Preparing the Critical Path to Acceptance of Toxicogenomic Data in Drug Safety Evaluation

❖Alex Nie, Johnson & Johnson PRD, 1000 Rt. 202 S, Raritan, NJ 08869, *anie@prdus.jnj.com*

Toxicogenomics is the application of genomics and bioinformatics in drug safety evaluation. On the one hand, the industry wants to find out safety problems for their drug candidates early to avoid higher developmental cost in later stages. On the other hand, the regulators want to encourage the use of toxicogenomics to prevent unsafe drugs from coming into clinical trials. What has been done so far in achieving a mutual understanding between the industry and the regulators for the use of toxicogenomics in drug safety evaluation? The application and acceptance of toxicogenomics have benefited greatly from a number of non-profit collaborative research programs such as the ILSI/HESI, HL7/CDISC/I3C, and the Critical Path Institute. This presentation will explain how these collaborative efforts helped to pave the way for acceptance of toxicogenomics data, and it will also show some examples from Johnson & Johnson's participation.

224 Resampling Methods in Small-Area Estimation ● ✪

Section on Survey Research Methods, Section on Government Statistics, Section on Nonparametric Statistics

Tuesday, July 31, 8:30 am–10:20 am

Jackknife and Bootstrap Methods for Small-Area Estimation: An Appraisal

❖J. N. K. Rao, Carleton University, School of Mathematics and Statistics, 1125 Colonel By Drive, Ottawa, ON K1S 5B6 Canada, *jrao@math.carleton.ca*

Key Words: Confidence intervals, Conditional properties, MSE estimation, Small area models

Jackknife and bootstrap methods have been proposed for small area estimation in recent years; in particular for mean squared error (MSE) estimation and confidence intervals. In this talk I will provide an overview and appraisal of the proposed methods. I will also present some results on a jackknife method for non-linear small area models that leads to area-specific MSE estimators.

On Small-Area Prediction Interval Problems

Snigdhansu Chatterjee, The University of Minnesota; ❖Partha Lahiri, University of Maryland, JPSM, 1218 Lefrak Hall, College Park, 20742, *plahiri@survey.umd.edu*; Huilin Li, University of Maryland

Key Words: Predictive distribution, Linear mixed model, Bootstrap, Coverage accuracy, Empirical Best Linear Unbiased Predictor

Empirical best linear unbiased prediction (EBLUP) method uses a linear mixed model in combining information from different sources of information. This method is particularly useful in small area problems. The variability of an EBLUP is measured by the mean squared prediction error (MSPE), and interval estimates are generally constructed using estimates of the MSPE. Such methods have shortcomings like undercoverage, excessive length and lack of interpretability. We propose a resampling driven approach. Simulation results demonstrate the superiority of our method over the existing ones.

225 Inference in Simple Dynamic Games with Multiple Equilibria ● ✪

JBES-Journal of Business and Economic Statistics

Tuesday, July 31, 8:30 am–10:20 am

The Identification Power of Equilibrium in Simple Games

❖Elie Tamer, Northwestern University, Department of Economics, 2001 Sheridan Drive, Evanston, IL 60208, *tamer@northwestern.edu*

Key Words: dynamic games, partial identificatoin, set inference, multiple equilibria

We explore the identification power that equilibrium assumptions provide. In particular, we explore the identified features of a game when we drop the assumption that the players are playing Nash equilibrium. Instead, we assume the players are rational. Rationality has been widely studied in the game theory literature, and in this paper, we adopt the rationalizability concept of Pearce and Bernheim. Here, we are able to examine the identification power of several layers of rational play in three games. In the first, we compare the identification power of rationality in bivariate discrete games with complete information to ones where equilibrium is imposed. We then examine games with incomplete information and characterize there the identification features of equilibrium. Finally, we examine the identification power in an independent private value auction.

226 Robust Inference and Model Selection

IMS, Biometrics Section, Section on Nonparametric Statistics

Tuesday, July 31, 8:30 am–10:20 am

Localized Model Selection by Cross-Validation

Key Words: model selection, localized model selection, cross validation

Traditional model selection theory and practice focus on selecting a single one that is deemed the best in a global sense. In high dimensional or complex data situations, when very distinct learning procedures are considered, it is often the case that the candidates perform differently in ranking in different regions. This motivates localized model selection that allows more flexibility to improve over the global model selection strategy. In this talk, we will focus on localized cross validation. Our goal is to select adaptively the procedure that performs the best locally among the candidate regression procedures. Theoretical and empirical results will be provided.

Selecting Models Using Incomplete Data

❖Gerda Claeskens, K.U. Leuven, ORSTAT, Naamsestraat 69, Leuven, International B-3000 Belgium, *gerda.claeskens@econ.kuleuven.be*

Key Words: model selection, missing data, AIC

We propose model selection methods which can deal with missing or incomplete data. More particularly, we focus on sets of data where the response variable is always observed, but some of the covariates are possibly missing. The relevant model selection question is which of the covariates should be included in the final model. We study the use of the EM algorithm, employing a method of weights, in combination with model selectors in the spirit of Akaike's information criterion AIC. In a first step the missing data mechanism is assumed to be ignorable. Extensions to other missingness mechanisms are possible. The model selection method is tested in a simulation study and illustrated by data analysis.

Predicting Future Responses Based on Possibly Misspecified Working Models

❖Lu Tian, Northwestern University, Dept. of Preventive Medicine, 680 N. Lake Shore Drive, Chicago, IL 60611, *lutian@northwestern.edu*; Tianxi Cai, Harvard University; Scott Solomon, Brigham and Women's Hospital; Lee-Jen Wei, Harvard University

Key Words: K-fold cross validation, Model misspecification, optimal prediction region

Under a general regression setting, we propose an optimal unconditional prediction procedure for future responses. The resulting prediction intervals or regions have a desirable average coverage level over a set of covariate vectors of interest. When the working model is not correctly specified, the traditional conditional prediction method is generally invalid. On the other hand, one can empirically calibrate the above unconditional procedure and also obtain its cross-validated counterpart. Various large and small sample properties of these unconditional methods are examined analytically and numerically. We find that the K-fold cross validated procedure performs exceptionally well even for cases with rather small sample sizes. The new proposals are illustrated with a real example.

227 Innovative Imputation Methods in Health Research ● ✪

Section on Health Policy Statistics, Biometrics Section, Section on Teaching Statistics in the Health Sciences, Section on Survey Research Methods, WNAR

Tuesday, July 31, 8:30 am–10:20 am

Using a Mixture Model for Multiple Imputation in the Presence of Outliers

❖ Michael R. Elliott, University of Michigan, M4041 SPHII, 1420 Washington Heights, Ann Arbor, MI 48109, *mrelliot@umich.edu*; Nicolas Stettler, Children's Hospital of Philadelphia

Key Words: latent class, survey sampling, MEMI, obesity, community health center

To obtain population-based inference in the presence of missing data and outliers, we develop a latent class model that assumes each observation belongs to one of K unobserved latent classes, with each latent class having a distinct covariance matrix. The latent class covariance matrix with the largest determinant is assumed to form an "outlier class," and we conduct inference after removing these outliers. As in Ghosh-Dastidar and Schafer (2003), we use multiple imputation to promulgate uncertainty in the outlier status. We extend their work by embedding the outlier class in a larger mixture model, consider penalized likelihood and posterior predictive distributions to assess model choice and fit, and construct the model to account for complex sample designs. We apply our methods to estimate obesity prevalence and body mass index (BMI) measures in the Healthy For Life Study.

Imputing Under-Reported Treatments Using Multiple Sources of Treatment Information in a Cancer Services Study

❖ Alan M. Zaslavsky, Harvard University, Health Care Policy, 180 Longwood Ave, Boston, MA 02115, *zaslavsk@hcp.med.harvard.edu*; Yulei He, Harvard Medical School

Key Words: imputation, multivariate probit, cancer, bayesian model, health services, measurement error

Cancer registry records, patient surveys, and administrative systems record adjuvant therapies (chemo and radiation) for cancer patients, but subject to underreporting that could bias analyses. We propose to impute true treatment status, using sample validation data from medical records, and analyze the imputed data. We extend an earlier study with a single outcome (provision of chemotherapy) and base data system (the registry), to multiple measures (provision of chemotherapy and radiation therapy) and multiple data systems (the registry, a patient survey, and Medicare claims). Bayesian hierarchical models for provision and reporting of multiple cancer therapies take into account their associations and multilevel structure, using related multivariate probit models for reporting of each therapy. The methodology is applied to data for patients with colorectal cancer in California.

Statistics and Lies: Correcting Questionnaire Ordering Effect via Multiple Imputation

❖ Xiao-li Meng, Harvard University, 1 Oxford Street 7th Floor, Dept of Statistics, Cambridge, MA 02138, *meng@stat.harvard.edu*; Jingchen Liu, Harvard University; Chih-nan Chen, Boston University; Margarita Alegria, Cambridge Health Alliance

Key Words: questionnaire ordering effect, survey design, Bayesian modelling, high-order interactions, psychiatric epidemiology, MCMC

National Latino and Asian American Study (NLAAS) is a complex survey of psychiatric epidemiology, with multiple embedded experiments with alternative question forms. One objective of multiple imputation is to create analytic datasets corrected for response biases due to defects of the survey instrument, such as increasing rates of negative responses over the course of the interview induced by the respondents' learning to use skip patterns to reduce interview time. The imputation modeling task is particularly challenging because of the complexity of the questionnaire, the small sample sizes for subgroups of interests, and the existence of high-order interactions among variables. This talk presents our Bayesian modeling and computation effort in addressing these problems.

228 The Bayesian Frequentist Interface: Is the Difference One of Methodology Or of Other Information? ✪

SSC, Section on Bayesian Statistical Science, ENAR, Section on Survey Research Methods

Tuesday, July 31, 8:30 am–10:20 am

Adaptive MCMC Using Stratified Proposals for Bayesian and Frequentist Inference

❖ Mylène Bédard, University of Warwick, Department of Statistics, Zeeman Building, Coventry, CV4 7AL United Kingdom, *M.Bedard@warwick.ac.uk*; Donald A.S. Fraser, University of Toronto; Augustive C.M. Wong, York University

Key Words: Metropolis-Hastings algorithm, Adaptive MCMC method, Frequentist vs. Bayesian, Regression example

In this talk, we study the equivalence between frequentist p-values and their corresponding Bayesian posterior survivor values (s-values), obtained when choosing a not too subjective prior. This is achieved by considering a simple regression example for which exact calculations can be obtained, and testing various hypotheses for the regression parameters. We then validate the obtained p-values and s-values using MCMC methods, which verify that they are uniformly distributed over (0,1) conditionally on some information provided by the data. We assess the numerical values for accuracy using the traditional Metropolis-Hastings algorithm, and we also use the asymptotic properties of our regression example to propose two different adaptive algorithms. These lead to a significant increase of the acceptance rate, meaning that the methods propose moves in a more wisely determined manner.

Improved Likelihood Inference for Vector Parameters with Discrete Data

❖ Nicola Sartori, University of Venice, San Giobbe, Cannaregio 873, Venezia, 30121 Italy, *sartori@unive.it*

Key Words: Contingency table, Discrete data, Likelihood asymptotics, P-value, Vector parameter

We consider the testing of compound hypotheses or vector component parameters in contingency tables and other discrete data models. In general models, asymptotic inference from the observed likelihood provides just first order accuracy. For the discrete data case Davison, Fraser and Reid (JRSSB, 2006) develop second order methods for assessing a scalar component parameter, which makes use of a data based reparameterization of exponential type. We extend this approach to obtain second order p-values for the assessment of compound or vector hypotheses. This makes use of marginal likelihoods and directional departure measures. The approach is illustrated on examples.

Default Priors for Frequentist and Bayesian Inference

❖Donald A.S. Fraser, University of Toronto, 100 St. George Street, Dept Statistics, Statsitics 6th floor, Toronto, ON M5S 3G3 Canada, *dfraser@utstat.toronto.edu*

Key Words: Bayesian, frequentist, default prior, Marginalization paradoxes

A location model has a natural default prior, the constant or flat prior in the location parameter. This gives a posterior interval which has standard coverage properties for a scalar parameter linear in the location coordinates but typically has biased coverage when it is nonlinear in the location components; this is the marginalization paradox of Dawid, Stone and Zidek (1973). For a general model with moderate continuity we develop second order flat priors for full and for component parameters; these provide second order coverage properties, respectively. This focuses the source of the marginalization paradoxes which is parameter curvature; and it shows how coverage bias can be avoided using priors targeted on the interest parameter.

229 Medallion Lecture III

IMS, General Methodology

Tuesday, July 31, 8:30 am–10:20 am

Functional Regression Analysis: Models, Methods, and Applications

❖Hans-Georg Müller, University of California, Davis, Department of Statistics, Davis, CA 95616, *mueller@wald.ucdavis.edu*

Key Words: Functional Data Analysis, Principal Component, Smoothing, Dimension Reduction

Functional regression analysis addresses the situation where predictors or responses in a regression setting include random functions. Early functional linear models were based on observing complete trajectories for the random functions. Generalized functional linear models can be used for classification. Recent extensions to be discussed include functional regression diagnostics, functional regression for sparse and noisy longitudinal data, time-varying functional models, functional additive and nonlinear regression, and the functional embedding of very high-dimensional data. A basic tool is the representation of random functions by functional principal component scores. Illustrations include data from longitudinal, demographic and gene expression studies. The presentation draws on joint work with Jeng-Min Chiou, Peter Hall, Xiaoyang Leng, Jane-Ling Wang, Ping-Shi Wu and Fang Yao.

230 The Future of Pro Bono Statistics ● ✪

Section on Statistical Consulting

Tuesday, July 31, 8:30 am–10:20 am

Volunteerism in the ASA

❖David Banks, Duke University, ISDS Box 90251, Durham, NC 27708, *banks@stat.duke.edu*

Key Words: volunteers, survey, Darfur, education

In 2005 the ASA established a Special Interest Group on Statistical Volunteerism. That group undertook pro bono projects in a number of areas, such as survey designs for Darfur, development of curricula for developing nations, and analyzing data from the Innocence Project. Not all of the efforts led to successful projects, but often things worked well. This talk describes the accomplishments, the obstacles, and the possibilities for the future.

A Role for Students in Pro Bono Statistics

❖Gayla R. Olbricht, Purdue University, Department of Statistics, STATCOM, 250 N. University Street, West Lafayette, IN 47907-2066, *ghobbs@stat.purdue.edu*

Key Words: pro bono statistics, student volunteerism, STATCOM, service-learning, statistical consulting, statistical consulting

In 2001, a graduate student at Purdue University founded Statistics in the Community (STATCOM). This voluntary graduate student-run program provides free statistical consulting to local governmental and nonprofit community groups. STATCOM fulfills a service-learning role by giving students the opportunity to serve society while applying classroom material to real-life situations. In 2006, STATCOM was awarded a Strategic Initiatives Grant from the ASA with the goal of serving as a resource to others interested in starting similar programs. One ultimate goal of these efforts is to establish a formal network among students engaged in pro bono statistics. This talk discusses how students can participate in pro bono statistics, the benefits of such student-driven volunteerism, achievements and lessons learned from the STATCOM Strategic Initiative, and ideas for the future.

Pro Bono Statistics and Public Policy

❖Yasmin H. Said, George Mason University, 4400 University Drive - Mail Stop 682, Department of Computational and Data Sciences, Fairfax, VA 22030, *ysaid99@hotmail.com*

Key Words: pro bono statistics, volunteerism, public policy, global warming, climate change, hockey stick

Efforts to persuade Congress to enact legislation that affects public policy are constantly being made by lobbyists who are paid by special interests. While this mode of operation is frequently extremely effective for achieving the goals of the special interest groups, it often does not serve the public interests in the best possible way. As counterpoint to this mode of operation, pro bono interaction with individual legislators and especially testimony in Congressional hearings can be remarkably effective in presenting a balanced picture. The debate on anthropogenic global warming has in many ways left scientific discourse and landed in political polemic. In this talk I will discuss our positive and negative experiences in formulating testimony on this topic.

231 Using Environmental Information for Rational Decisionmaking ● ✪

Section on Statistics and the Environment, Section on Risk Analysis

Tuesday, July 31, 8:30 am–10:20 am

Virus and Indicator Occurrence in Ground Water Sources of Public Drinking Water

❖Michael J. Messner, U.S. Environmental Protection Agency, Office of Ground Water and Drinking Water, 1200 Pennsylvania AVE, NW (4607M), Washington, DC 20460, *messner.michael@epa.gov*; Crystal Rodgers, U.S. Environmental Protection Agency

Key Words: environmental, hierarchical

Promulgated in 2006, the Ground Water Rule will require utilities to test their source waters for fecal contamination whenever total coliforms are found to be present in their distribution systems. When fecal contamination is found in a well, action must be taken to eliminate or reduce the associated public health risk. To assess the health and economic impacts of these requirements, EPA needed to assess the occurrence of enteroviruses and E. coli (an indicator of fecal contamination) in ground water wells. This paper describes the occurrence data that were used to inform a hierarchical model of virus and indicator occurrence. Results are presented in terms of parameter estimates, baseline risk, and probabilities of taking action as a result of the new rule.

A Bayesian Approach to EPA's Data Quality Objectives Process

❖Paul K. Black, Neptune and Company, Inc., 8550 W. 14th Street, Suite 100, Lakewood, CO 80215, *pblack@neptuneinc.org*; Mark Fitzgerald, Neptune and Company, Inc.; Tom Stockton, Neptune and Company, Inc.

Key Words: DQOs, Bayesian methods, Scientific Method, statistical planning

In order to improve the quality of its decisions, the Environmental Protection Agency recommends the use of the Data Quality Objectives (DQO) Process. The intent of the DQO process is to offer a systematic approach to planning for data collection consistent with the goals of the Scientific Method. EPA has focused on the use of classical statistical methods to implement the implied decision analysis. Arguably, the Scientific Method is best implemented using Bayesian methods. When the DQO process was first introduced, computer programs for implementing a Bayesian approach were not available. With the advent of new algorithms for Bayesian analysis, this is no longer an issue. We will present a contrast between the two approaches, demonstrate the benefits of using a Bayesian approach to DQOs and, more generally, to decisionmaking, and provide some real world examples of its implementation.

Strengthening Environmental Decisionmaking Through Data Quality Objectives

❖John Warren, U.S. Environmental Protection Agency, 1200 Pennsylvania Ave NW, (2811R), Washington, DC 20460, *warren.john@epa.gov*

Key Words: Decision making, Data Quality Objectives

In order to improve the quality of its decisions, the Environmental Protection Agency demands the use of systematic planning for data collection and the generation of appropriate performance or acceptance criteria. In order to achieve this goal, the Agency recommends the use of the Data Quality Objectives (DQO) Process, a logical seven-step, iterative procedure rooted in the scientific method. By making the process easy to follow and yet sufficiently rigorous, the Agency has successfully strengthened its decision making capabilities.

232 Multiple Comparisons and Experimental Replication ● ✪

Section on Bayesian Statistical Science

Tuesday, July 31, 8:30 am–10:20 am

Synthesis of Bayes Factors and Estimation for Large Samples: Adaptive Hypothesis Testing Intervals (AHTI)

❖Luis R. Pericchi, University of Puerto Rico, San Juan, P.O. Box 23355, Department of Mathematics, San Juan, PR 00931-3355, *luarpr@gmail.com*; Maria E. Perez, University of Puerto Rico, San Juan

Key Words: Adaptive Hypothesis Testing Intervals(ahti), Synthesis, Bayes Factors, Estimation, Hypothesis Testing, Objective Bayes

As a first Synthesis, we propose a calibration of p-values and significance levels that convey specific guidelines in how to diminish the alpha-levels as the sample sizes grows, to alleviate the discrepancy of Bayes Factors and Classical Testing (assuming it is accepted that type I error is not fixed but can diminish as evidence accumulates). This decrease in alpha-levels is appropriate for testing with huge sample sizes. There is a second Synthesis involved very important to unify Bayesian model selection procedures, the Synthesis between Probability Intervals and Bayes Factors: A pervasive idea in the whole of Statistics is to reject null hypotheses when point nulls are outside intervals. This is wrong (from a Bayesian point of view) when confidence levels are held fixed.

Why There Are So Many Contradicted or Exaggerated Findings in Highly Cited Clinical Research

❖Ying Yuan, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Dept of Biostatistics Unit 447, Houston, TX 77054, *yyuan@mdanderson.org*; Valen Johnson, The University of Texas M.D. Anderson Cancer Center

Key Words: test statistics, Bayesian model, p value

It is common for reports of clinical research regarding the effectiveness of new interventions to be contradicted in subsequent trials. This is particularly disturbing when high-impact research is involved. We analyzed 42 highly-cited (>1000 citations) controlled randomized clinical studies published between 1990 and 2003 and previously analyzed in Ioannidis (2005). Seven (18.5%) out of 38 positive studies were contradicted or were found to have overstated effects in subsequent studies. To understand this fact, we extracted the test statistics from the original articles and analyzed their values using a simple Bayesian model. Based on this model, we relate the reported p-values from these studies to posterior probabilities of associated hypotheses. Conclusions from this study are somewhat surprising and have important implications for consumers of classical testing methods.

Bayesian Decision Theory for Multiplicities

❖Kenneth Rice, University of Washington, Box 357232, F-600 Health Sciences Building, 1705 NE Pacific Street, Dept of Biostatistic, Seattle, WA 98195-7232, *kenrice@u.washington.edu*

Key Words: Utility, Loss function, Interval Estimation, Decision Theory, Multiple Comparisons

Many Bayesian analyses conclude with a summary of the posterior distribution, thus summarizing uncertainty about parameters of interest. This is not enough, as it neglects to state what it is about the parameters that we want actually want to know. Formally, deciding our criteria for a 'good' answer defines a loss function, or utility, and is usually only considered for point estimation problems. For interval estimation, we provide sensible, interpretable loss functions which formally justify some 'standard' but essentially ad-hoc Bayesian intervals. Developing these measures of utility for problems with multiple parameters is straightforward, and the Bayes rules remain attractive and simple. Direct connections can be made with frequentist methods of multiplicity-adjustment.

233 Radiation Epidemiology and Statistical Methods ● ✪

Section on Statistics in Epidemiology, ENAR

Tuesday, July 31, 8:30 am–10:20 am

Health Effects of Induced Radiation from the Atomic Bombings of Hiroshima and Nagasaki

❖Fumiyoshi Kasagi, Radiation Effects Research Foundation, Hijiyama Park 5 2, Minami Ward, Hiroshima, 732-0815 Japan, *kasagi@rerf.or.jp*

Key Words: Atomic bomb, Neutron, Induced radiation, Health effects

Few studies have been conducted on health risk from neutron-induced radiation. The Radiation Effects Research Foundation's (RERF's) has undertaken a study using their Life Span Study (LSS) cohort to assess the health effects of induced radiation from the atomic bombings of Hiroshima and Nagasaki. The LSS cohort contains a sub-cohort of people that were not in the city (NCI) at the time of the bomb and were not directly exposed to atomic bomb radiation. This paper provides the first assessment of the potential risk from induced radiation using the data from the RERF NIC cohort. Risk assessments for entrants on the first day (which are assumed to have the highest risk) are compared to entrants on the second, and third days. Comparison are made accounting for differences on a number of baseline variables.

Modeling Effects of Radiation Exposure on Disease Outcomes with Radiation-Influenced Risk Factors (RFs), Using Adult Health Study (AHS) Data

❖Kenneth Kopecky, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N Mailstop M4C102, PO Box 19024, Seattle, WA 98109-1024, *kkopecky@fhcrc.org*; Lori A. Williams, University of Washington

Key Words: Radiation epidemiology, Time-dependent confounding, Cardiovascular mortality, Radiation Effects Research Foundation

The AHS is a longitudinal study of radiation effects in survivors of the A-bombings of Hiroshima and Nagasaki. The cohort, initially about 20,000 survivors, has been offered medical examinations since 1958, with about 17,000 examined at least once, averaging 11 exams in 23 exam cycles to 2004. Linking AHS data to disease outcomes may allow testing if apparently radiogenic disease risks arise in part from radiation-related changes in disease RFs. For example risks of hypertension (HT) and of death from heart disease and stroke (together cardiovascular morality, CVM) increase with radiation dose. Since HT is a RF for CVM mortality, radiation's apparent effect on CVM might be due in part to radiation-related HT. Methods for estimating the radiation effect on disease risk that account for time-dependent confounding using RF data from the AHS will be discussed with the example of CVM and HT.

Biodosimetry in the A-Bomb Survivors' Cohort

❖Daniel Stram, University of Southern California, 1540 Alcazar Street, Room 220G, Lost Angeles, CA 91001, *stram@usc.edu*

Key Words: Radiation, Measurement Errors, Instrumental Variables

The follow-up of the A-bomb survivors is a primary source of information concerning the human health effects of exposure to radiation. Quantification of risk is based upon a physical dosimetry system that estimates dose without reference to any biological endpoints. There are a large number of survivors for whom data on one or more of the following endpoints exists: (1) acute radiation symptoms immediately following exposure; (2) chromosome aberrations and mutational analysis for the GPA locus; (3) occurrence of radiation induced cataracts and; (4) ESR-based dose measurements of tooth enamel. We describe the use of these biological dosimetry data in order to estimate and give confidence intervals for parameters in the dosimetry error model. We also outline a proposal to develop a "Biologically enhanced" dosimetry system for primary risk estimation for the follow-up study.

Uncertainty of Estimated Radiation Doses Among Atomic Bomb Survivors: Using Multiple Sources of Information on Errors in the Main Exposure Variable of Health Effects Studies

❖Harry Cullings, Radiation Effects Research Foundation, 5 2 Hijiyama Park Minami ku, Hiroshima, 732-0815 Japan, *hcull@rerf.or.jp*; Eric J. Grant, Radiation Effects Research Foundation; Sachiyo Funamoto, Radiation Effects Research Foundation; Sachiko Teranishi, Radiation Effects Research Foundation

Key Words: exposure uncertainty, radiation dosimetry, biodosimetry, dose error

Studies of atomic bomb survivors in Hiroshima and Nagasaki use doses estimated by a dosimetry system as the main exposure variable; these inherit uncertainty from many sources. There is no subset of survivors with dose estimates suitable for validation, but there are many sources of partial information about dose uncertainty. Progress has been made in the past with methods for dealing with dose error based on assumed joint distributions of true and estimated doses. This talk will focus on sources of ancillary information that could improve uncertainty estimates, and their relation to error models. Examples include reliability of recall in re-interviewed survivors, population density, spatial proximity in families, influence of distance and shielding uncertainty based on physical propagation of error, biodosimetry, environmental measurements, and data on acute signs of radiation exposure.

Descriptive Models for Excess Risks Following Radiation Exposure

❖Dale Preston, Hirosoft International, 1335 H St, Eureka, CA 95501-2331, *preston@hirosoft.net*

Key Words: radiation effects, dose response, hazard functions, excess rates

Log-linear models are commonly used to describe relative risks in epidemiology. Such models are of limited use in describing dose response relationships. These limitations include the implicit non-linearity of the dose response and difficulties in interpreting effect modifiers that describe effects relative to baseline risks. Modeling the excess relative risk (ERR=1 minus the relative risk) as the product of a dose response function and effect modifiers such as gender or attained age leads to useful descriptions of the risk. Models of this form can also easily be used to describe excess rates. ERR and excess rate models provide contrasting interpretations of the excess risk that can lead to insights into and hypotheses about the nature of exposure effects. I will describe risk models used in radiation effects studies with examples based on the atomic bomb survivor data.

234 New Developments in Causal Inference ● ✪

Biometrics Section, ENAR

Tuesday, July 31, 8:30 am–10:20 am

Structural Nested Mean Models for Time-Varying Causal Effect Moderation

❖Daniel Almirall, University of Michigan, 650 Hidden Valley Club Drive, Apartment 118, Ann Arbor, MI 48104, *dalmiral@umich.edu*; Thomas R. Tenhave, University of Pennsylvania; Susan Murphy, University of Michigan

Key Words: Causal effect moderation, Structural nested mean model, G-estimator, 2-stage regression, bias-variance trade-off, time-varying covariates

This talk considers the problem of assessing effect moderation in longitudinal settings in which treatment is time-varying and so are the covariates said to moderate its effect. The main challenges of assessing causal effect moderation in the time-varying setting are discussed. Intermediate Causal Effects that describe time-varying causal effects of treatment conditional on past covariate history are introduced and considered as part of Robins' Structural Nested Mean Model. Two estimators of the causal effects are presented: The first is a 2-Stage Regression Estimator, which can be used with standard regression software. The second is Robins' G-Estimator. The methodology is illustrated using longitudinal data from the PROSPECT study. Our goal is to estimate the effects of time-varying adherence to the intervention conditional on time-varying covariates that may modify these effects.

Estimation and Inference for the Causal Effect of Receiving Treatment on a Multinomial Outcome

❖Jing Cheng, University of Florida, 1329 SW 16th Street, Suite 5130, Gainesville, FL 32610, *jcheng@biostat.ufl.edu*

Key Words: Randomized trials, noncompliance, multinomial outcomes, causal effect, bootstrap

This paper considers the analysis of two-arm randomized trials with noncompliance which have multinomial outcomes. We define and estimate the compliers average causal effect (CACE) in these trials with the likelihood method. A simulation study shows that the likelihood ratio statistic for testing the CACE does not follow a chi-squared distribution. We propose a bootstrap version of a likelihood ratio test for the CACE in these trials. The methods are illustrated by an analysis of data from a randomized trial of an encouragement intervention to improve adherence to prescribed depression treatments among depressed elderly patients in primary care practices.

Statistical Analysis of Randomized Experiments with Nonignorable Missing Outcomes

❖Kosuke Imai, Princeton University, Department of Politics, Princeton, NJ 08544, *kimai@princeton.edu*

Key Words: causal inference, noncompliance, instrumental variables, average treatment effects, sensitivity analysis, identification

Missing data are frequently encountered in the statistical analysis of randomized experiments. In this paper, I propose statistical methods that can be used to analyze randomized experiments with a nonignorable missing binary outcome where the missing-data mechanism may depend on the unobserved values of the outcome variable itself. I first introduce new identification strategies for the average treatment effects and complier average causal effects. I then derive the maximum likelihood estimator and its asymptotic properties, and discuss possible estimation methods. Since the proposed identification assumption is not directly verifiable from the data, I show how to conduct a sensitivity analysis based on the parameterization that links the key identification assumption with the causal quantities of interest. I apply the proposed methods to analyze data from two randomized experiments.

Correcting for Survivor Treatment-Selection Bias with a Structural Failure-Time Model: Survival in Oscar Award–Winning Performers

❖Xu Han, University of Pennsylvania, 400 Jon M Huntsman 3730 Walnut Street, Philadelphia, PA 19104, *hanxu3@wharton.upenn.edu*; Dylan Small, University of Pennsylvania; Vishal Patel, University of Pennsylvania; Dean Foster, University of Pennsylvania

Key Words: survival analysis, causal inference, structural failure time model, G-estimation

We study the causal effect of winning an Oscar Award on an actor or actresses' survival. Does the increase in stature from a performer winning an Oscar increase the performer's life expectancy? Previous studies of this issue have suffered from survivor treatment selection bias, that is, candidates will have more chance to win Oscar Awards if they live longer, and winning Oscar Awards at a certain age is also an indicator of health status. To correct this bias, we adapt Robins' rank preserving structural failure time model and G-estimation method. We show in simulation studies that our approach corrects the survivor treatment selection bias contained in other approaches and apply our model to the Oscar data.

Estimating Surrogate Endpoints Defined by Principal Causal Effects

❖Shiro Tanaka, University of Tokyo, 731 Hongo Bunkyo, Tokyo, 113-0033 Japan, *shiro@epistat.m.u-tokyo.ac.jp*; Yutaka Matsuyama, University of Tokyo; Yasuo Ohashi, University of Tokyo

Key Words: advanced prostate cancer, monotonicity assumption, principal stratification, surrogate endpoint

In clinical trials, comparison of treatments for the outcome of primary interest such as survival time may require a long follow-up before yielding useful results. Because of this, there is increasing interest in the use of surrogate endpoints to make decisions about treatment efficacy. Frangakis and Rubin (2002) proposed a principal surrogate, which was defined by potential values of surrogate endpoints. However, it is impossible to observe joint distribution of potential surrogate endpoints without untestable assumptions. In this talk, we propose an estimation method of the principal surrogate under monotonicity assumption, which states that there is no subject who would not have a response under test arm, but have a response under control arm. We also applied the proposed method to a clinical trial data of advanced prostate cancer.

235 Statistical Analysis of Medical Device Studies ● ✪

Biopharmaceutical Section, ENAR, WNAR

Tuesday, July 31, 8:30 am–10:20 am

Assessing Poolability of Data for a Clinical Trial: Why the Data Should Be Poolable and Why the Data Should Not Be Poolable

❖David Naftel, The University of Alabama at Birmingham, LHRB 790, 1530 3rd Ave South, Birmingham, AL 35294-0007, *dnaftel@uab.edu*

Key Words: Interactions in a randomized clinical trial, Pooling of data, Subgroup analysis

Clinical trials require pooling data from multiple sites to obtain adequate sample size. The FDA requires that such pooling be justified. Components of pooling are discussed in this paper with examples. Was the proposed study design and conduct the same at each site? These components are nonnegotiable. Are the sites similar as to the baseline variables? The procedure is to compare baseline variables across sites. Lack of differences may indicate a homogenous sample that does not represent the target patients. Multiple tests almost guarantee that sites will be found different. We recommend assessing overlap among sites to understand similarities rather than differences. Do differences exist in outcomes? We recommend identifying the statistical explanation of the differences. Nonpoolable data allows investigation of interactions and may provide a representative study.

Primary Endpoint Sensitivity Analysis on Missing Data in Medical Device Trials

❖Gina Garding, Boston Scientific Corporation, 4100 Hamline Avenue N., South Town Square, MS 9-315, St. Paul, MN 55112, *gina.garding@bsci.com*; Nathan Carter, Boston Scientific Corporation; Eric Bass, Boston Scientific Corporation; Paul Cernohous, Boston Scientific Corporation

Key Words: Sensitivity Analysis, Missing Data, Data Imputation, Medical Device Trials

It is becoming increasingly common that the FDA requests sensitivity analyses to assess the effect missing data could have on conclusions drawn from a medical device clinical trial. The assumption that the data is missing at random is sometimes questionable, requiring a sensitivity analysis. This analysis may be based on hypothetical situations, including "case" analyses (best, average and worst) or failure-threshold analysis. This presentation will compare the results of these analyses for binary endpoints.

Challenges and Issues of Intent-to-Treat Analysis in Cardiac Device Studies

❖Li Ding, Medtronic Inc., 1015 Gramsie Road, MS Z230, Shoreview, MN 55126, *li.ding@medtronic.com*; Steve Boeh, Medtronic Inc.

Key Words: Intent-to-Treat, Cardiac Device, Simulation, Cardiac Resynchronization Therapy, Implantable Cardiac Defibrillator

Intention-to-treat analysis (ITT) is an important consideration in medical device and pharmaceutical studies. Medical devices differ fundamentally from pharmaceutical drugs and biological products in their mechanisms of action. This nature of differences contributes to more challenges and issues of ITT analysis in device trials. In this research presentation, the common issues and deviations during the course of randomized controlled cardiac device studies will be studied. Their impact on the analysis results will be investigated through the simulation study.

Assessment of the Success of the Propensity Score Estimation and Appropriate Adjustment in a Drug Eluting Stent Post-Marketing Registry

❖Yun Lu, Boston Scientific Corporation, 100 Boston Scientific Way, Marlborough, MA 01752, *luy@bsci.com*; Kevin Najarian, Boston Scientific Corporation; Aijun Song, Boston Scientific Corporation

Key Words: Propensity Score, Stepwise Logistic Regression, Subclassification, F test, Drug Eluting Stent

Propensity scores are used frequently in observational studies to adjust for covariate imbalance and reduce bias regarding group comparisons. However, the distribution of propensity scores generated by simple logistic regression may not overlap enough for valid group comparisons. To assess the success of the propensity score estimation and adjustment, covariates, interactions and quadratic terms are included in a logistic regression model with stepwise selection. The distribution of the propensity score is then subclassified into quintiles. F-tests from a sequence of models constructed by gradual refinement of the initial model are used to examine the improvement of the balance across subclasses. Box plots provide visual evaluation of the distributions among groups. This method is applied to a post-marketing drug eluting stent registry study.

Bayesian Approach to Meta-Analysis in Medical Device Trials Using Mixed Effects Models

❖Lijuan Deng, Boston Scientific Corporation, 100 Boston Scientific Way, M1 Dept of Biostatistics Clinical Science, Marlborough, MA 01752, *dengl@bsci.com*; Hong Wang, Boston Scientific Corporation; Liang Li, Genzyme Corporation

Key Words: Bayesian approach, meta-analysis, Monte Carlo Markov chain, medical device trial, mixed effects models

A mixed effects model allowing for study-level variability is proposed to incorporate treatment effects from different medical device trials. This model accounts for variability at both patient-level and study-level. The patient-level treatment effects are assumed to have their own distributions within each study, while the study-level treatment effect follows a common underlying distribution across studies. In this Bayesian meta-analysis model, the posterior distributions and credible intervals of treatment effect are estimated using Markov Chain Monte Carlo method with WinBUGS, by adjusting for other baseline covariates such as lesion length, vessel diameter and diabetic status.

236 Nonparametric Methods for Complex and Difficult Data ✪

IMS, Section on Nonparametric Statistics

Tuesday, July 31, 8:30 am–10:20 am

Covariance Selection by Sparse Regression

❖Jie Peng, University of California, Davis, Department of Statistics, One Shields Ave, Davis, CA 95616, *jie@wald.ucdavis.edu*; Pei Wang, Fred Hutchinson Cancer Research Center

Key Words: covariance selection, sparse regression, sparsity, high dimension

In this paper, we propose a joint sparse regression approach for covariance selection under the setting of p>n. This method depends on the overall sparsity of the concentration matrix, but does not make assumptions on neighborhood sparsity of individual variables. We study the performance of this new approach under various simulation settings. We also apply the method on high-dimensional array data for genetic network inference where identification of hubs (genes with many connections) is of great interests. We demonstrate that our method is more powerful in hub identification compared to existing methods. We also show that by taking into account of the sparsity of the concentration matrix, a better estimation of the covariance matrix can be obtained.

On Single-Index Models

❖Yan Yu, University of Cincinnati, Cincinnati, OH, *yan.yu@uc.edu*; Zhou Wu, University of Cincinnati; Haiqun Lin, Yale University

Key Words: roughness penalty, multivariate nonparametric estimation, penalized spline, local linear, quantile regression

Single-index models are an important tool for multivariate nonparametric estimation. By reducing the dimensionality from that of a general covariate vector to a univariate index, single-index models avoid the so-called "curse of dimensionality." In this talk, firstly I will give a brief overview of single-index models. Secondly I will introduce single-index varying coefficient models for both regression and dependent data, where we focus on the estimation, inference, and forecasting of the single-index coefficient models under dependence by a penalized spline approach (P-spline). Finally I will discuss some recent results for single-index conditional quantiles using local linear estimation. Both real data applications and simulation studies will be presented.

CLUES: A Nonparametric Clustering Method Based on Local Shrinking

❖Xiaogang Wang, York University, Department of Math Stat, 4700 Keele Street, Toronto, ON M3J 1P3 Canada, *stevenw@mathstat.yorku.ca*

Key Words: Clustering, Local Shrinking, K-nearest neighbor, Number of clusters,

The authors propose a novel non-parametric clustering method based on non-parametric local shrinking. Each data point is transformed in such a way that it moves a specific distance toward a cluster center. The direction and the associated size of each movement are determined by the median of its K-nearest neighbors. The optimal value of the number of neighbors is determined by optimizing some commonly used index functions that measure the strengths of clusters generated by the algorithm. The number of clusters and the final partition are determined automatically without any input parameter except the stopping rule for convergence. The experiments on simulated and real datasets suggest that the proposed algorithm achieves relatively high accuracies when compared with classical clustering algorithms.

Longitudinal Modeling When Response and Time-Dependent Covariates Are Measured at Different Time Points

❖Joel Dubin, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 2H2 Canada, *jdubin@uwaterloo.ca*; Xiaoqin Xiong, University of Waterloo

Key Words: longitudinal, smoothing, lagged association

In this talk, we will discuss some flexible methods to handle both association and temporal sequencing of distinct longitudinal measures, where the measures may be of mixed type (e.g., one continuous, the other binary) and recorded on nonuniform grids and different time points from one another. A smoothing step will be involved. The methods will be demonstrated on a dataset of hemodialysis patients, where longitudinal measures of health outcomes (e.g., infection) were recorded at different time points than longitudinal physiologic measures such as serum C-reactive protein levels (a marker for inflammation). An interesting scientific question to answer is whether experiences of infection follow or predate inflammation.

Regularized Autoregressive Approximation: A Bridge Between Estimation and Model Selection

❖Bei Chen, University of Waterloo, 12 McDougall Road, Waterloo, ON N2L2W5 Canada, *beiaries@gmail.com*; Yulia R. Gel, University of Waterloo

Key Words: time series analysis, estimation, regularization, autoregressive models, model order selection, shrinkage regression

One of the common approaches in time series is to approximate the true model by "long" autoregressive (AR) models. Usually the order of an AR model is selected by information criterions and then parameters are obtained by the least squares method etc. However, as sample size increases, it often implies that the model order should be refined, and hence all the parameters need to be recalculated. Thus, the cost of approximation increases. In this talk we propose a regularized approximation which enables to estimate AR coefficients with different level of accuracy. Thus, the model selection step is "smoothed" as the sample size increases. The regularized AR approximation can be considered therefore as a version of a shrinkage regression. We illustrate applications of the regularized AR approximation to estimation of ARMA models, frequency detection and forecasting of long memory processes.

237 Bayesian Applications in Marketing ● ✪

Section on Bayesian Statistical Science

Tuesday, July 31, 8:30 am–10:20 am

A Simultaneous Quantile Regression Model for Customer Linkage Analysis

❖Jeffrey Dotson, The Ohio State University, 2100 Neil Avenue, Columbus, OH 43210, *dotson_83@cob.osu.edu*; Joseph J. Retzer, Maritz Research; Greg M. Allenby, The Ohio State University

Key Words: Services Marketing, Simultaneity, Quantile Regression

Quantile regression for a simultaneous system of equations is developed to study the relationship between customer and employee satisfaction. Customers interact with many employees, and employees serve many customers, such that a one-to-one mapping between customers and employees is not possible. Analysis proceeds by relating the distributions of customer and employee responses. Such analysis is commonly encountered in marketing when data are from independently collected samples. Our model facilitates analysis among quantiles of the distributions, allowing comparison among better-performing and worse-performing units. We demonstrate our model in the context of retail banking, where drivers of customer and employee satisfaction are shown to be quantile-dependent.

Disentangling Preferences, Inertia, and Learning in Brand Choice Models

❖Sangwoo Shin, University of Rochester, Simon Graduate School of Business, 10 Manhattan Square Drive, Rochester, NY 14607, *shins2@simon.rochester.edu*; Sanjog Misra, University of Rochester; Dan Horsky, University of Rochester

Key Words: Preferences, Learning, Choice, Heterogeneity, Scanner, RJMCMC

Using a unique dataset that contains stated preferences and actual purchase data for the same group of consumers we attempt to untangle the effects of preference heterogeneity and state dependence, and to determine the exact nature of the latter. We propose a hierarchical model in which

consumers are heterogeneous in the order of the brand choice process as well as in their preferences and responsiveness to marketing mix. The proposed model is designed to encompass three different types of consumer experience based behavior: zero-order, inertia, and learning. We apply a Reversible Jump MCMC sampling scheme to sample across component processes and a Metropolis-Hastings/Gibbs step within each component process. Our results suggest that the extent of state dependence/preference heterogeneity is spuriously over/underestimated in the absence of preference information.

A Model of State Dependence Based on Auto- and Cross-Correlated Exponential Processes with an Application to Household Panel Data

❖ Xiuyun Zhang, The Ohio State University, 641 Stark CT, Columbus, OH 43210, *xyzhang@stat.ohio-state.edu*; Thomas Otter, The Ohio State University; Mario Peruggia, The Ohio State University

Key Words: state dependence, brand choice, purchase timing, point process model, auto-correlation, cross-correlation

We use an epidemic-type point process model to describe state-dependence, i.e. the impact of past purchases on future purchases, in a novel way. We look at household panel data that records purchase timing, quantity and brand purchased as well as marketing mix variables such as price and promotional activity for one product category. Brand choice and purchase timing are jointly dependent variables conditional on past purchases, purchase quantity, and marketing mix variables. In our model the baseline rate is expressed as a function of time-varying marketing mix variables (price, display and feature) and is brand specific. We consider both the impact of past purchases of the same brand (auto-correlated effect) and of past purchases of other brands (cross-correlated effect) on future brand choice and purchase timing. Inference is based on MCMC techniques.

Bayesian Semiparametric Estimation of Random Coefficients Discrete Choice Models Using Aggregate Data

❖ Sylvie Tchumtchoua, 210 Quinebaug, Storrs, CT 06269, *tchumtchoua@yahoo.fr*

Key Words: Order-based dependent Dirichlet Process, discrete choice models, consumer learning, state dependence, unobserved heterogeneity, advertising

In the literature on estimating random coefficients discrete choice models with aggregate data, typically price is the only strategic variable to firms, the model does not incorporate any form of dynamics in brand choice, and consumer heterogeneity is modeled using a unimodal parametric distribution. We adapt Bayesian semiparametric methods to a structural random coefficients discrete choice model with advertising goodwill effect, learning, state dependence, and unobserved heterogeneity. By semiparametric we mean consumer heterogeneity is modeled using the recent order-based dependent Dirichlet Process suggested by Griffin and Steel; this approach helps recover a richer variety of heterogeneity distributions while allowing the nonparametric distribution to depend on consumers' covariates. The model is applied to analyze demand for ready-to-eat breakfast cereal in selected U.S. cities.

238 Leonard J. Savage Awards ● ✪

Section on Bayesian Statistical Science

Tuesday, July 31, 8:30 am–10:20 am

Semiparametric Models in Bayesian Event History Analysis Using Beta Processes

❖ Pierpaolo De Blasi, University of Turin, Piazza Arbarello 8, Turin, 10122 Italy, *pierpaolo.deblasi@unito.it*

Key Words: Bayesian semiparametrics, Bernstein-von Mises theorem, beta processes, competing risks model, hazard regression, Poisson random measures

We consider two topics of event history analysis: proportional hazards regression and competing risks models. Both are studied within new semiparametric formulations and the aim is to develop a full Bayesian treatment. The hazard regression model we consider is a variant of the Cox model, as it postulates a logistic relative risk function. The Bayesian construction involves a beta process for the cumulative baseline hazard and a Jeffreys-type density for the regression coefficients. The posterior distribution is derived and a Bernstein-von Mises (BvM) theorem is reached for our class of priors. Competing risks data describe failure times with multiple endpoints. We model cause-specific hazards (CSH) via the conditional probability of a failure type and the overall hazard. We propose a beta process for the overall cumulative hazard and derive a BvM theorem for the posterior of the CSHs.

Density Estimation and Dimension Reduction with Logistic Gaussian Process Priors

❖ Surya Tokdar, Carnegie Mellon University, 5000 Forbes Ave, Department of Statistics, Pittsburgh, PA 15213, *stokdar@stat.cmu.edu*

Key Words: Markov chain Monte carlo, L-1 convergence, Smoothing, Bayesian semiparametric modeling

Logistic Gaussian process priors provide an extremely flexible framework to build smooth, nonparametric models for density estimation. Lack of efficient computing methods, however, have kept these priors somewhat underused and understudied. We propose a new method that allows efficient computation with these priors for density estimation. This method can be easily generalized to other applications with models based on transforms of a Gaussian process. We illustrate this with an application to sufficient dimension reduction in multiple regression. In either application we prove that strong posterior consistency properties obtain for a rich class of models.

Structure and Sparsity in High-Dimensional Multivariate Analysis

❖ Carlos Carvalho, Duke University, 832 Marilee Glen Ct, Durham, NC 27705, *carlos@stat.duke.edu*

Key Words: Gaussian Graphical Models, Sparse Factor Models

As scientific problems grow in terms of both expanding parameter dimension and sample sizes, structure and sparsity become central concepts in practical data analysis and inference. By allowing complex high-dimensional problems to be modeled through low-dimensional underlying relationships, sparsity helps to simplify estimation, reduce computational burden and facilitate interpretation of large scale datasets. This talk addresses the issue of sparsity modeling primarily in the context of Gaussian graphical models and sparse factor models. To convey the main ideas of this work the talk will focus on the extension of conditional independence ideas from Gaussian graphical models to multivariate dynamic linear models. After presenting the development of this new class of models I will describe applications of such models in large financial time series and portfolio allocation problems.

Bayesian Treed Gaussian Process Models

❖ Robert Gramacy, University of Cambridge, Wilberforce Rd, Cambridge, CB3 0WB United Kingdom, *bobby@statslab.cam.ac.uk*

Key Words: recursive partitioning, nonstationary spatial model, nonparametric regression, Bayesian model averaging, sequential design, computer simulator

Computer experiments often require dense sweeps over input parameters to obtain a qualitative understanding of their response. However, such sweeps are unnecessary in regions where the response is easily predicted; well-chosen designs could allow a mapping of the response with far fewer simulation runs. Thus, there is a need for computationally inexpensive surrogate models and an accompanying method for selecting small designs. I explore a semiparametric nonstationary modeling methodology for addressing this need that couples stationary Gaussian processes with treed partitioning. A Bayesian perspective yields an explicit measure of (nonstationary) predictive uncertainty that can be used to guide sampling. The benefit of adaptive sampling is illustrated through several examples, including a motivating example which involves the computational fluid dynamics of a NASA re-entry vehicle.

239 Opportunities and Challenges of Integrated Datasets To Serve Policy Research: The Medicaid Undercount Study ● ✪

Section on Government Statistics, Section on Health Policy Statistics, Social Statistics Section

Tuesday, July 31, 8:30 am–10:20 am

Opportunities and Challenges Facing the Construction of Integrated Datasets To Serve Policy Research Purposes

❖ Michael Davern, The University of Minnesota, 2221 University Ave SE, Suite 345, Minneapolis, MN 55414, *daver004@umn.edu*

Key Words: administrative data, survey data, policy research, survey methods

Survey data have limitations for policy research yet they continue to be widely used because, despite these limitations surveys are the only sour source for some data we need to predict, evaluate and learn about the impact of potential changes and actual changes in policy. Public program administrative data that are used to keep track of program beneficiaries are beginning to be used in combination with survey data to create new linked data products. The potential uses of these linked data products for improving policy research are tremendous but many issues need to be dealt with before these linked data products begin to overtake the popularity of survey public use microdata for policy analysis. In this paper we set out a research agenda for improving linked data files for policy research.

Analysis of Medicaid Enrollment Status in the Current Population Survey

❖ Ming-Yi K. Mah, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *mingyi.k.mah@census.gov*; Dean M. Resnick, U.S. Census Bureau

Key Words: Medicaid, Undercount, Currrent Population Survey

Medicaid recipient estimates from surveys are lower than administrative data counts of enrollment. It is believed that Current Population Survey (CPS) non-respondents have a higher percentage of Medicaid enrollees than the respondents of the survey, and thus would account for a portion of the underestimate. This paper compares the CPS 2002 response rates for people enrolled and not enrolled in Medicaid. Medicaid data were matched with the CPS respondent and non-respondent data to compare the rates of Medicaid enrollment. The results show that the CPS non-respondents do not have a higher percentage of Medicaid enrollees than the respondents, and that some demographic groups are in fact more likely to respond to the CPS if they are receiving Medicaid benefits.

How Do Surveys Differ in Reporting the Quality of Reported Medicaid Enrollment Data: NHIS, CPS, and State Surveys

❖ Kathleen Call, The University of Minnesota, 420 Delaware St SE MMC729, Minneapolis, MN 55455, *callx001@umn.edu*

Key Words: measurement error, surveys, administrative data, Medicaid undercount, uninsurance

Survey measurement error can contribute to the tendency for surveys to show lower counts of Medicaid enrollment than administrative data (i.e., Medicaid Undercount). This happens two ways: first, respondents can misreport that they are uninsured when they really have Medicaid, biasing survey estimates of the uninsured and those on Medicaid. Second, respondents can answer they have coverage other than Medicaid, biasing the Medicaid count and not the uninsured count. The first form of error is more detrimental to policy research than the second. We compare the prevalence of both error scenarios using three types of surveys in which we know the Medicaid enrollment status of respondents: four state surveys, the Current Population Survey, and the National Health Interview Survey. Preliminary analysis shows that rates of the first type of error are much higher in the CPS than other surveys.

Larger Statistical and Policy Implications of Moving Toward Interagency Collaborations and Integrated Data

❖ Sally Obenski, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *sally.m.obenski@census.gov*

Key Words: linked data, Medicaid, undercount, adjustments

This paper discusses the broader implications of using integrated data sets to improve statistics and better inform policy makers. It compares and contrasts standard approaches to statistical research v the approach took in the Medicaid Undercount Study, focusing on the benefits as well as the challenges. Additionally, it discusses the need for a mechanism that inserts results such as those from the Medicaid Undercount Study into published estimates from the CPS. For example, repeating the analysis on a yearly basis could provide both the Census Bureau and policy makers with potential "adjustments" or footnotes to published estimates on Medicaid enrollment.

240 Alternative Approaches to the Introductory Applied Statistics Course

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 8:30 am–10:20 am

Alternate Topic Sequencing in the Introductory Applied Statistics Course

❖ John Gabrosek, Grand Valley State University, 1 Campus Drive, Dept of Statistics, Allendale, MI 49401, *gabrosej@gvsu.edu*

Key Words: statistics pedagogy, statistics education, statistical inference

The statistical background of students graduating from institutions of higher learning is often limited to a single course in introductory applied statistics. It is imperative that this course adequately prepare students to become knowledgeable consumers of information. Students leaving the course need to have a firm grasp of the role that statistical inference plays in decision-making. Unfortunately, the traditional sequencing of the introductory applied statistics course places the important topic of statistical inference at the end of the course. In this talk we discuss the results of a study to teach introductory statistics "sequencing by statistical case" rather than statistical topic.

Integrating Data Analysis with Parametric Inference in Undergraduate Statistics

❖Larry Weldon, Simon Fraser University, Statistics and Actuarial Science, 8888 University Drive, Burnaby, BC V5A 1S6 Canada, *weldon@sfu.ca*

Key Words: graphics, data analysis, resampling, software, undergraduate, curriculum

A rising interest in graphical methods and resampling seems at odds with traditional parametric inference. The role of graphs is often portrayed as a supplementary technique for checking parametric assumptions, yet it deserves to be recognized as an independent method of analysis. The problem for the curriculum designer is to construct an appropriate merger of the two approaches for undergraduate courses. This session encourages discussion about ways to do this effectively.

Comparison of a Traditionally Taught Introductory Statistics Course to a Completely Online Version of the Same Course

❖ Gail Tudor, Husson College, One College Circle, Bangor, ME 04401, *tudorg@husson.edu*

Key Words: Online vs traditional learning, attitudes toward statistics

Student performance and attitudes were compared between an online introductory statistics course (n=10) and the same course taught face-to-face (n=30). Online materials consisted of slides with corresponding audio and notes, practice quizzes, and discussions. The courses had the same professor, textbook and covered the same materials. There was no significant difference in mean exam scores or course grades. Regarding attitudes toward statistics, the online class did not change over the semester, but the traditional section experienced an 11% increase in positive attitudes. Regarding the professor's interaction, enthusiasm, and knowledge, all students reported positively, but the traditional class was more likely to strongly agree. The online class was slightly happier with the pace of the course, but wanted more group work. Past online student evaluations support these findings.

Team-Based Learning in an Introductory Applied Statistics Course

❖ Paul Fields, Brigham Young University, 223A TMCB, Provo, UT 84602, *pjfields@byu.edu*

Key Words: Introductory Statistics, Team-Based Learning, Engineering, Quality, Learning Outcomes

An innovative approach was implemented in an introductory applied statistics course directed primarily toward undergraduate engineering students. The course is unique in two ways: First, although the course covers all of the usual introductory statistics topics, the material is presented in the engineering context of improving processes and products. Second, the course is structured entirely in a team-based learning format. All learning activities and assessments are conducted in student teams similar to how engineers actually work professionally. The implicit and explicit connections to the "real-world" increase students' interest and engagement in the course. Students demonstrate greater mastery of the course material both during and after the course.

Active Student Learning Model of Statistics Education

❖ Ralph Turner, University of the Sciences in Philadelphia, 600 S 43rd Street, Philadelphia, PA 19104, *r.turner@usip.edu*

Key Words: Statistics, Education, Undergraduate, Active-Learning

The Active Student Learning Model of statistics education is based on two core assumptions. Embedding statistics training in the context research design enhances students' comprehension of statistics and experiential work causes students to actively engage in statistical reasoning and thinking leading to greater knowledge and usable skills. Class time is structured for activities, aided by Just-in-Time Lectures. Students create Note Sets and Exams based on readings, use statistical software to conduct analyses of data, conduct an experiment involving all aspects of the process including literature review, proposal writing, power analysis, IRB review, database management, data cleaning, statistical analyses, manuscript preparation and formal presentation. Grades are based on students' demonstration of growth in statistical reasoning, and thinking skills.

241 Biostatistics and Bioinformatics at the Interface ● ✪

WNAR, Biometrics Section, ENAR, Section on Health Policy Statistics

Tuesday, July 31, 8:30 am–10:20 am

How Well Does Structural Equation Modeling (SEM) Reveal Weakened Brain Anatomical Connections? An fMRI Simulation Study

❖ Jieun Kim, National Institutes of Health, 4710 Bethesda Ave APT1511, Bethesda, MD 20814, *kimjieun@nidcd.nih.gov*; Barry Horwitz, National Institutes of Health

Key Words: Structural Equation Modeling, path coefficients, effective connectivity, functional MRI, delayed match-to-sample task

SEM, a well-known statistical technique for brain effective connectivity analysis in functional MRI studies, is often used to study patients. This study investigates whether a damaged brain anatomical connectivity can be revealed by SEM. We applied realistic fMRI generated by a large-scale neural network performing a visual delayed match-to-sample task to SEM. We simulated 20 normal subjects and 20 patients for whom the strength of anatomical connection from inferior temporal (IT) to prefrontal cortex (PFC) was reduced to 20% of its normal value. We found that path coefficients for the IT/PFC connection and all feedback pathways from PFC to other regions were significantly reduced in patients. Thus, although a weakened effective connection can be attributed to a weakened anatomical connection, a reduced effective connection may not always reflect a weakened anatomical connection.

Correlation Analysis of the Transcriptomic and Proteomic Data in Platelet

❖ Chen Ji, Stony Brook University, 69 Eastwood Blvd, Centereach, NY 11720, *cji@ic.sunysb.edu*; Wei Zhu, State University of New York at Stony Brook

Key Words: canonical correlation, resampling, platelet

Conventional wisdom implies a close relationship between mRNAs and their corresponding proteins. In this paper, based on a human platelet study, we examined three different approaches to calculate the correlation. They are the Pearson correlation, the Spearman correlation, and the canonical correlation. We also applied the bootstrap resampling approach without normality assumption and obtained a significant correlation within the unique platelet system.

Optimal Designs Accounting for Potentially Missing Observation in Quantal Dose Responses

❖Inyoung Baek, Novartis Pharmaceuticals, Fabrikstrasse 4, 5th floor, Basel, 4056 Switzerland, *inyoung.baek@novartis.com*

Key Words: Logit model, Bayesian optimal criterion, Information matrix, non-missing probability

Dose-response experiments are routinely conducted to examine response rates at various dose levels of interest. Due to side effects, lack of efficacy or other reasons, not all responses are observed in practice. Consequently, our final design can be inefficient. Baek et al. (2006) proposed to construct single and multiple-objective optimal designs for a binary dose response study when potentially missing observations are incorporated at the onset of the study. In this paper, we compared the performance between commonly used designs in clinical trials and the proposed optimal designs accounting for potential missing observations in terms of the precision of related parameter estimations, for example, the estimation of certain dose levels or the underlying dose response model parameters.

Power and Sample Size in Genetic Association Studies in Presence of Genotyping Error

❖Wonkuk Kim, Stony Brook University, Dept of Applied Math and Statistics, Stony Brook, NY 11794-3600, *wkim@ams.sunysb.edu*; Qing Wang, Stony Brook University; Derek Gordon, Rutgers University; Stephen J. Finch, Stony Brook University

Key Words: Likelihood Ratio Test, Mixture Model, Case/Contol Study, Noncentrality Parameter, SNP

The test of whether the distribution of genotypes of a single nucleotide polymorphism (SNP) in a control population is the same as the distribution in an affected population can be made out using the 2x3 test of independence. When the genotyping is determined by an underlying continuous measure that is the mixture of three normal components, the likelihood ratio test (LRT) of the equality of mixing proportions is an alternative. We compare the performance of these tests by first calculating the power of the LRT and the relative efficiency of the 2x3 test to the LRT. When the minor SNP allele frequency is less than 0.2 in both cases and controls and the separation between genotype components is small, the LRT is more efficient than the 2x3 test. We present detailed tables of efficiencies and the limiting behavior of the relative efficiency.

Locally Optimal Weighting and Classifier Selection in Ensembles

❖Melissa Fazzari, Stony Brook University, 24 Forest Drive, Centerport, NY 11721, *mjfazzari@yahoo.com*; Hongshik Ahn, Stony Brook University; Wei Zhu, State University of New York at Stony Brook

Key Words: ensembles, classification, bias-variance decomposition, nearest neighbors

Ensemble-based classification improves generalization accuracy through a mix of variance and bias reduction. The largest gains are found through the aggregation of strong, but diverse classifiers. For each individual test point, the best set of classifiers and their ensemble weights may be highly varied. We examine locally optimal classification, a weighting scheme for combining classifiers based on local performance. Weights are determined based on a bootstrap estimate of variance for each classifier at each training point. Classification of test instances is achieved by combining the predictions across classifiers using the weights of the training set nearest-neighbors. Other methods of ensemble selection and combination are also explored, including low bias combining and a correlation-based approach. Variable importance across all ensemble members is examined and summarized.

242 Statistical Graphics for Everyday Use? ● ✪

Section on Statistical Graphics, Section on Statistical Consulting

Tuesday, July 31, 8:30 am–10:20 am

Statistical Graphics for Everyday Use?

❖John Emerson, Yale University, 438 Humphrey St, New Haven, CT 06511, *john.emerson@yale.edu*; ❖Frederick Wicklin, SAS Institute Inc., SAS Campus Drive, S3030, Cary, NC 27511, *rick.wicklin@sas.com*; ❖Leland Wilkinson, SPSS Inc., 233 South Wacker Drive 11th fl, Chicago, IL 60606, *leland@spss.com*

There is a discrepancy between research in statistical graphics, their implementation in statistical software packages, and their use in daily data analysis applications. This session will discuss various modern graphical methods and strategies in graphical data analysis with regard to their usefulness and generality. Whereas open-source software allows prototyping and adopts new research ideas fast, it often fails to establish standards in graphical methods. Commercial software is more conservative in adding graphical methods, but brings them to a diverse user base. In this context, several questions should be discussed: Which commonly available methods are only "chart-junk," and which are outdated? Are there general rules for the construction of "good graphics"? How should missing values and survey weights be treated? Will more recent methods become standard graphical tools?

243 Analysis of Censoring/Truncated Data ● ✪

Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

Nonparametric Maximum Likelihood Estimation of the Incidence Rate Using Data from a Prevalent Cohort Study with Follow-Up

❖Vittorio Addona, Macalester College, 1600 Grand Avenue, Saint Paul, MN 55105, *addona@macalester.edu*; Masoud Asgharian, McGill University; David Wolfson, McGill University

Key Words: prevalent cohort, right censoring, left truncation, incidence rate, nonparametric maximum likelihood estimator

Prevalent cohort studies with follow-up often require fewer resources than incident cohort studies. We discuss nonparametric maximum likelihood estimation of a constant incidence rate that uses the well-known incidence-prevalence relationship. Efficient estimation of the incidence rate using this relationship, and based only on onset and failure/censoring times collected from a prevalent cohort study with follow-up, poses some major difficulties. In such studies, the onset times of prevalent cases are ascertained and subjects with shorter survival times are less likely to be recruited. We show how it is possible to adjust for the onset times of these "missing" cases when estimating the incidence rate. We apply our approach to data from the Canadian Study of Health and Aging to estimate age-specific incidence rates of dementia amongst elderly Canadians.

Semiparametric Analysis for Recurrent Event Data with Time-Dependent Covariates and Informative Censoring

❖ Chiung-Yu Huang, National Institutes of Health, 6700A Rockledge Drive, Room 5234, Bethesda, MD 20817, *huangchi@niaid.nih.gov*; Jing Qin, National Institutes of Health; Mei-Cheng Wang, Johns Hopkins University

Key Words: Comparable recurrence times, Frailty, Informative censoring, Pairwise pseudolikelihood, Proportional rate model

We present a semiparametric model of recurrent event data that allows censoring time and recurrent event process to be correlated via frailty. This flexible framework includes both time-dependent and time-independent covariates, while leaving the distributions of frailty and censoring time unspecified. We estimate the effect of time-dependent covariates by constructing a pseudo likelihood based on comparable pairs of event times. For the estimation of the baseline cumulative rate function and the regression coefficients of time-independent covariates, we derive a modified product-limit estimator with bias correction in risk sets, and solve estimating equations formulated based on expected number of observed recurrent events. Numerical studies demonstrate that the proposed methodology performs well for practical sample sizes.

Nonparametric Estimation in a Markov 'Illness-Death' Process from Interval-Censored Observations with Missing Intermediate Transition Status

Halina Frydman, New York University; ❖ Michael Szarek, Pfizer Inc., 47 West Devonia Avenue, Mount Vernon, NY 10552, *mszarek99@yahoo.com*

Key Words: illness-death model, interval censoring, missing transition status

We develop a nonparametric method for estimation of the cumulative transition intensities in an irreversible time-nonhomogeneous three-state Markov chain, or 'illness-death' model, from intermittent observations of the process. The observations result in interval-censored transition times to the intermediate state and, for a subset of individuals, unknown status prior to death or right censoring. Existing nonparametric methods for estimation of the transition intensities in an illness-death model with interval-censored intermediate transition times rely on the assumption that the state of an individual is always known prior to death or prior to right censoring. Our approach extends these methods to the situation where the state of an individual is unknown in an interval of time for a subset of subjects due to periodic assessment of the intermediate event.

Inference for Three-State Progressive Disease Models with Common Periodic Observations

❖ Beth Ann Griffin, RAND Corporation, 1200 S Hayes Street, Arlington, VA 22202, *bethg@rand.org*; Stephen Lagakos, Harvard School of Public Health

Key Words: Arm-in-cage experiments, Progressive Disease models, Interval censoring, Time-to-event data

We develop statistical methods for designing and analyzing arm-in-cage experiments used to test the efficacy of insect repellents. Efficacy of a repellent can be described using a progressive three-state model in which the first two states represent varying degrees of protection (no landing and landing without biting) and the third state occurs once protection is completely lost (biting). In the experiments, subjects place their arms into cages periodically to observe whether mosquitoes land and/or bite. Since subjects within a treatment group follow the same cage visit schedule, transition times between states are interval censored into one of several fixed intervals. We develop an approach for estimating the parameters of interest when sojourn times are dependent and study design considerations for these experiments. The proposed methods are illustrated on real and simulated data.

Semiparametric Estimation with Recurrent Event Data Under an Informative Monitoring Period

❖ Akim Adekpedjou, University of South Carolina, 216 LeConte Building, Department of Statistics, Columbia, SC 29208, *adek@stat.sc.edu*; Edsel A. Pena, University of South Carolina

Key Words: Informative censoring, Koziol-Green model, Martingale, Semiparametric estimation, Efficiency

We consider a biomedical study which monitors the occurrences of a recurrent event for n subjects over a random observation window for each subject. We assume that the distribution of the random observation window is informative regarding the distribution of event time. The problem of semiparametric estimation of the cumulative hazard and consequently of the gap-time is considered under a model of informative censoring called the Koziol-Green model. We derive a Nelson-Aalen and Kaplan-Meier type estimators respectively for the cumulative hazard and the gap-time distribution function under the specified model. Asymptotic and small sample properties of the proposed estimators are established. The proposed estimators are compared to the nonparametric estimator of the proposed in Pena et al. (2001, JASA) to ascertain any efficiency gain achieved by exploiting the Koziol-Green structure.

Comparing in vivo Tumor Growth with Response-Dependent Right-Censoring

❖ Wenquan Wang, The University of Alabama at Birmingham, 1824 6th Ave S, NP 2540, Birmingham, AL 35294, *wwang@uab.edu*

Key Words: tumor growth, in vivo, response-dependent censor, area under curve

Tumor xenografts study has an important role in assessment of anti-tumor agents. Traditionally, tumor growth and animal survival are the main approaches in assessment for subcutaneous tumor xenografts. For intra-peritoneal tumor xenografts, survival was the main approach for assessment. With rapid growth in small animal imaging techniques, tumor growth becomes as important as survival in assessment. However, comparing tumor growth between treatment groups is challenging when there are heterogeneous response-dependent right-censoring, such as animal dies due to rapidly growing large tumor or animal is sacrificed due to large tumor. Linear random effects with censoring probability, conditional linear model, and some non-parametric or semi-parametric approaches have been proposed in the past. A simple and straightforward test statistic based on area under growth curve is proposed.

Evaluation of a Binary Response Variable as a Surrogate for Survival: A Gynecologic Oncology Group Study

❖ Virginia Filiaci, GOG Statistical & Data Center, Elm and Carlton Streets, Buffalo, NY 14263, *gfiliaci@gogstats.org*; Randolph L. Carter, University at Buffalo-SUNY

Key Words: Surrogate endpoint, survival, indirect effect, explained variance, clinical trial

Identification of valid surrogates for survival that lead to a decrease in study times or the resources necessary to complete clinical trials is important. Methods were developed to evaluate a binary response variable as a surrogate for survival using data from a single trial. Two criteria are suggested for perfect surrogacy: 1) the total effect of treatment on survival must be comprised solely of its indirect effect through the surrogate; and 2) survival must be perfectly predicted by the surrogate and treatment. Two statistics are proposed in terms of these criteria. The first is the proportion of the total treatment effect on mean censored survival time truncated at

T0 that is mediated through the surrogate. A measure defined by Korn and Simon is adopted as the second. These methods will be applied to data from a randomized trial of 263 women with metastatic endometrial cancer.

244 Bayesian Statistical Methodology and Its Application to Biometrics

Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

A Bayesian Approach to a Logistic Regression Model with Incomplete Information

❖Taeryon Choi, University of Maryland, Baltimore County, 1000 Hilltop circle, Baltimore, MD 21250, *tchoi@math.umbc.edu*; Mark J. Schervish, Carnegie Mellon University; Ketra Schmitt, Battelle Memorial Institute; Mitchell Small, Carnegie Mellon University

Key Words: Aggregate information, Dose-response study, Logistic regression, MCMC, WinBUGS

We consider a set of independent Bernoulli trials with possibly different success probabilities that depend on covariate values. However, the available data consist only of aggregate numbers of successes amongst subsets of the trials along with all of the covariate values. We still wish to estimate the parameters of a modeled relationship between the covariates and the success probabilities (e.g., a linear logistic regression model). In this paper, estimation of the parameters is made from a Bayesian perspective by using a Markov Chain Monte Carlo (MCMC) algorithm based only on the available data. The proposed methodology is applied to both simulation studies and real data from a dose-response study of a toxic chemical, perchlorate.

A Bayesian Method for the Detection of Epistasis in Quantitative Trait Loci

❖Edward Boone, Virginia Commonwealth University, Department of SSOR, 1001 W Main St, Richmond, VA 23284, *elboone@vcu.edu*

Key Words: Bayesian Statistics, Quantitative Trait Loci, Markov Chain Monte Carlo Model Composition, Genetics

Epistasis or the interaction between loci on a genome is of great interest to geneticists. However, this problem is complicated by the actual definition of interaction and the number of loci and interactions versus the number of observations. Often researchers find there simply isn't enough data to sort through a large number of loci on a genome and then compound that with the interaction effects. This presentation discusses the definition of interaction in Recombinant Inbred Lines (RIL) and proposes a method to search for the loci as well as the interaction (epistasis) effects. This Bayesian method utilizes a Markov Chain Monte Carlo Model Composition approach using restricted sample spaces. The method is illustrated with the Arabidopsis thaliana plant considering cotyledon opening against 38 loci using 158 lines.

Group Sequential Designs and Analyses for Clinical Trials Based on a Utility Setting

❖Yi Cheng, Indiana University South Bend, 1700 Mishawaka Ave, South Bend, 46634, *ycheng@iusb.edu*

Key Words: Bayesian method, Clinical trial, Conditional expectation, Group sequential, Error rates, Utility

A Bayesian group sequential design is proposed for a comparative two-armed clinical trial using utility. Utility functions are set to address multiple issues: to treat patients in clinical trials more effectively and to learn about the better arms more quickly. Data are cumulated using Bayesian method. At each interim analysis, the decision to terminate or to continue the trial is based on the conditional expected utility. The performance of the proposed design is studied in terms of Bayesian optimality, as well as the frequentist error rates for regulatory settings.

Statistical Model for Detecting Multiple eQTLs

❖Wei Zhang, Harvard University, 1 Oxford St, Statistics Dept, Cambridge, MA 02138, *weizhang@fas.harvard.edu*; Jun S. Liu, Harvard University; Jun Zhu, Rosetta Inpharmatics, LLC

Key Words: eQTL, gene module, epistasis, pleiotropy, MCMC

Treating mRNA transcript abundances as quantitative traits and mapping gene expression quantitative trait loci for these traits has been studied in many organisms. Due to the large number of gene expression values and genetic markers, it is still a challenging question to researchers where these associations are and how eQTLs affect expression levels. We will present a statistical model to describe the associations between gene expression and genetic markers. Unlike tradition eQTL analysis, our method treats genes with similar expression values and linked with similar markers as a module. It searches for the module and its linked markers simultaneously. The linkage is defined in a statistical way such that marker interactions are automatically considered. Simulation studies and real data examples will be presented to illustrate the method.

Empirical Bayes Estimators in Stratified Random Sampling/Ranked Set Sampling

❖Tetsuji Ohyama, Kurume University, 67 Asahimachi, Biostatistics Center, Kurume City, International 830-0011 Japan, *tetsuji-ohyama@cec.med.kurume-u.ac.jp*; Jimmy A. Doi, California Polytechnic State University; Takashi Yanagawa, Kurume University

Key Words: biased estimator, infinite population, U statistics, relative efficiency

In this paper, we discuss the estimation of population characteristics in a stratified random sampling in an infinite population framework and consider a use of prior information by empirical Bayes method. Underlying distribution is assumed unknown, and U-statistics is employed for constructing the estimator incorporated with the prior values of the characteristics. The optimal empirical Bayes estimator is obtained, but it contains unknown parameters and those parameters must be replaced in practice. Simulation is conducted to show the gains in efficiency of the proposed estimator, showing the gain is 1.1~1.5 times larger than the unbiased estimator in terms of the relative efficiency if prior values are close to the true value.

Linear Model Selection Based on Estimation of Model Bias

❖Andrew Neath, Southern Illinois University Edwardsville, Department of Mathematics, Edwardsville, IL 62026, *aneath@siue.edu*; Letitia Downen, Southern Illinois University Edwardsville; Joseph Cavanaugh, The University of Iowa

Key Words: model selection, Gauss discrepancy, Mallows' Cp

Model selection criteria often arise by constructing estimators of measures known as expected overall discrepancies. The expected overall Gauss (error sum of squares) discrepancy for a fitted linear model can be decomposed into a term representing the error due to estimation of unknown parameters and a term representing the approximation error, or bias, due to model misspecification. Since estimation error depends only on model dimension, a known quantity, the selection problem reduces to the problem of estimating the bias term for each fitted candidate model. In this talk, we consider various estimators, both frequentist and Bayesian, of bias

for fitted linear models and consider how best to quantify the uncertainty inherent to a model selection problem. The most well known selection criterion within this framework is Mallows' Cp.

Empirical Bayes Methods for Controlling the False Discovery Rate with Dependent Data

❖Weihua Tang, Bristol-Myers Squibb Company, 506 Dillion Court, North Brunswick, NJ 08902, *weihua.tang@bms.com*; Cun-Hui Zhang, Rutgers University

Key Words: multiple testing, false discovery rate, empirical bayes, dependent data

False discovery rate (FDR) has been widely used as an error measure in large scale multiple testing problems, but most research in the area has been focused on procedures for controlling the FDR based on independent test statistics or the properties of such procedures for test statistics with certain types of stochastic dependence. Based on an approach proposed in Tang and Zhang (2005), we further develop Empirical Bayes methods for controlling the FDR with dependent data. We implement our methodology in a time series model and report the results of a simulation study which demonstrate the advantages of proposed Empirical Bayes approach. This is joint work with Cun-Hui Zhang.

245 Sequence Analysis and Genetic Association Studies

Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

Mining ChIP Data for Transcription Factor Binding Sites by the Addition of Background Sequences

❖Yu Liang, National Institute of Environmental Health Sciences, 111 T W Alexander Drive, Research Triangle Park, NC 27709, *liangy3@niehs.nih.gov*; Leping Li, National Institute of Environmental Health Sciences

Key Words: Transcriptional factor binding sites, Position weight matrices, False selection rate, ChIP experiments

Identifying the true transcriptional factor binding sites is a essential step for understanding the gene regulation. In this paper, we proposed a new approach to identify transcriptional factor binding sites in ChIP datasets when the number of sequences with motif is unknown. Our method is designed to control the false selection rate (FSR), defined as the proportion of sequences without motifs included in the selected sequence sets, in the meantime optimize the PWM. To characterize the performance of our method, we added a known number of background sequences to the real ChIP dataset. A motif discovery procedure is applied to the mixed dataset. The PWM is updated through monitoring the proportion of background sequences falsely selected.

Complex Demodulation Applied to bZIP and bHLH-PAS Protein Domains

❖Charles Smith, North Carolina State University, Department of Statistics, 513E Cox Hall, Raleigh, NC 27695-8203, *bmasmith@stat.ncsu.edu*; Zhi Wang, North Carolina State University; William R. Atchley, North Carolina State University

Key Words: periodicity, entropy, functional region

Proteins are built with molecular modular building blocks such as an alpha-helix, beta-sheet, loop region and other structures. This is an economical way of constructing complex molecules. Periodicity analysis of protein sequences has allowed us to obtain meaningful information concerning their structure, function and evolution. Here complex demodulation (CDM) is introduced to detect functional regions in protein sequence data. More specifically, we analyzed bZIP and bHLH-PAS protein domains. Complex demodulation provided insightful information about changing amplitudes of periodic components in protein sequences. It was found that the local amplitude minimum or maximum of the 3.6-aa periodic component is associated with protein structural or functional information since the extrema are mainly located in the boundary area of two structural or functional regions.

A Probabilistic Approach for Phylogenetic Sequence Selection

❖Michelle Lacey, Tulane University, 6823 St Charles Avenue, Department of Mathematics, New Orleans, LA 70118, *mlacey1@tulane.edu*

Key Words: phylogeny, molecular sequences, quartets

The objective of many phylogenetic studies is to reconstruct the evolutionary history of one or more specific genes, and in these cases similarity search engines are often employed to identify closely related molecular sequences from a variety of species. Such searches return hundreds of candidate sequences for a given study, and choosing among these can be an arduous task. However, the importance of thoughtful sequence selection has been largely overlooked in the development of practical phylogenetic tools. Through the derivation of bounds for the error probability associated with the reconstruction of four-leaf trees, known as quartets, as a function of the quartet topology and sequence lengths, we present the foundations for an algorithm that will facilitate phylogenetic analyses by specifying those quartets most likely to be correctly reconstructed from a set of aligned sequences.

Variable Selection in Logistic Models for Testing SNP-SNP Interactions

❖Huiyi Lin, The University of Alabama at Birmingham, WTI 153, 1530 3rd Ave S, Birmingham, AL 35294-3300, *hylin@uab.edu*; Renee Desmond, The University of Alabama at Birmingham; Yung-Hsin Liu, Pharmaceutical Product Development Inc. ; Seng-jaw Soong, The University of Alabama at Birmingham

Key Words: single nucleotide polymorphism, gene interaction, logistic model, variable selection

Complex disease traits are associated with single nucleotide polymorphisms (SNPs) interactions. Logistic models are commonly used for variable selection in testing SNP-SNP interactions. No study has examined evidence-based variable selection for high order interactions in logistic models. This simulation study was designed to compare variable selection procedures in logistic models for testing SNP-SNP interactions. Data on 10 SNPs were simulated for 400 and 1000 subjects (case/control ratio=1). The simulated model included one main effect and two 2-way interactions. The variable selection procedures included automatic selection (stepwise, forward and backward), 2-step selection, AIC-based and BIC-based selection. The hierarchical rule effect was examined. We reported stepwise selection without hierarchical rule was preferred based on the percentages of true and false positive terms.

Likelihood-Based Procedures for Disease Gene Localization with General Pedigree Data

❖Shuyan Wan, Merck & Co., Inc., 126 E Lincoln Ave, RY 80M 162, Rahway, NJ 07065, *shuyan_wan@merck.com*; Shili Lin, The Ohio State University

Key Words: disease gene localization, confidence interval method, model averaging, importance sampling, linkage analysis

We proposed a confidence interval method for disease gene localization by testing every position on each chromosome of interest for its possibility of being a disease locus and including those not rejected into the interval. Two generalized likelihood ratio tests with or without model averaging (GLRT/MA and GLRT) were proposed to perform the test. Null distribution based on GLRT was estimated by importance sampling method. We also proposed asymptotic approaches based on both GLRT and GLRT/MA as alternatives that are much more efficient computationally but depends on the reliability of the limiting distributions. Besides its efficiency, the asymptotic procedure based on GLRT/MA also takes model uncertainty into consideration. Performance of various methods was compared by ROC-like curves based on both simulated data and a real data set from Genetic Analysis Workshop.

Statistical Methods for Genetic Association Studies Based on Haplotype Sharing

❖Glen Satten, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mailstop K23, Atlanta, GA 30341, *gas0@cdc.gov*; Andrew Allen, Duke University

Key Words: haplotype, association, genetics, case-parent trio, family data, case-control study

Haplotype phase structure may help discover genetic variants associated with disease. However, haplotype-based methods become complex with many loci. Haplotype sharing statistics have been proposed as a solution. These methods consider haplotype similarity instead of modeling each haplotype because, if a recent disease-causing mutation is introduced once into a population, then haplotypes with the mutation should be similar to each other. Thus, one compares the extent of similarity between case and control haplotypes (for case-control data) or transmitted and untransmitted haplotypes (for case-parent data). Existing haplotype sharing statistics use permutation, making them unsuitable for whole genome studies or simulation experiments. We give a class of haplotype sharing statistics and derive its asymptotic properties. A Crohn's disease data set illustrates our approach.

Constructing Gene Association Network for Rheumatoid Arthritis Using the Backward Genotype-Trait Association (BGTA) Algorithm

❖Yuejing Ding, Columbia University, 1255 Amsterdam Ave, Room 1018, New York, NY 10027, *yd2122@columbia.edu*; Tian Zheng, Columbia University; Iuliana Ionita-Laza, Harvard University; Lei Cong, Columbia University; Shaw-Hwa Lo, Columbia University

Key Words: BGTA, GTD score, Gene-gene interaction, Gene Association Network

In this paper, we propose a two-stage analysis of whole-genome scan data using the BGTA algorithm, and also demonstrate how to construct gene association network based on the results from the BGTA algorithm. We will illustrate this method using a genome scan SNP dataset ascertained by North American Rheumatoid Arthritis Consortium (NARAC). This method is able to capture both the marginal and interaction effects of multiple susceptibility loci. This novel gene association network provides, for the first time in the literature, the possible interaction between markers/genes which may be useful for the biological interpretation.

246 Issues with Survival Data in Clinical Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

Comparison of Methods for Handling Early Dropouts in Analysis of Event Rates

❖Yuting Zhang, Amgen Inc., 1120 Veterans Blvd., ASF21-2, South San Francisco, CA 94080, *ytkzhang@gmail.com*

Key Words: Dropouts, Event rates

To compare event rates or subject incidence rates in different treatment groups is a common analysis in medical research with fixed treatment duration. The issue of early dropouts is a big challenge to obtain the unbiased estimates of event rate. The widely used worse case analysis (treating dropout as event) or responder analysis (ignoring the dropouts) may result in underestimating or overestimating the event rates. In this paper, we compared different methods, such as the worst case imputation methods, Kaplan-Meier estimation, multiple imputations, etc. to handle the early dropouts through a simulation study. The different dropout mechanism was also considered.

Conditional Power for Stratified Logrank Tests

❖Chung-Kuei Chang, Cephalon, Inc., 18 Moores Rd, Malvern, PA 19355, *schang@cephalon.com*; Lothar Tremmel, Cephaloh, Inc.

Key Words: conditional power, logrank test, stratification, survival analysis, interim analysis

To make a go/no-go decision at interim analyses in confirmatory clinical trials, it is important to predict the chance of success at the final analysis based on the interim data and various assumptions in the future data. Lan and Wittes (1988) proposed the calculation of the conditional power based on the B-value theory, in which the expectation and variance of the conditional Z statistic at the final analysis were derived. However, their theory may not apply to stratified analyses. The issue is how to weight each stratum appropriately in order to estimate the variance of the pooled Z statistic. In the analysis of time-to-event data using the stratified logrank test, we provide a precise derivation to calculate the conditional power based on the decomposition of the logrank statistic proposed by Tsiatis (1981).

Optimal Cost-Effective Designs of Proof of Concept Trials

❖Cong Chen, Merck & Co., Inc., 1019 Brassington Dr, Collegeville, PA 19426, *cong_chen@merck.com*

Key Words: efficiency score, oncology, optimal design

This presentation discusses optimal cost-effective designs for proof of concept (PoC) trials. Unlike a confirmatory registration trial, a PoC trial is exploratory in nature and sponsors of such trials have the liberty to choose type I error rate and the power. The decision is largely driven by the perceived probability of having a truly active treatment per patient exposure (a surrogate measure to development cost), which is naturally captured in an efficiency score to be presented. Optimization of the score function leads to type I error rate and power (and therefore sample size) for the trial that is most cost-effective. The idea is applied to derive optimal trial-level, program-level and franchise-level design strategies. From the examples provided herein, the audience should be able to gain much-needed insight in similar design problems and participate in strategic decisions.

Nonparametric Confidence Intervals for Survival Probabilities

❖ Kun Chen, Bristol-Myers Squibb Company, 5 Research Parkway, Wallingford, CT 06492, *kun.chen@bms.com*

Key Words: Kaplan Meier product-limit method, confidence interval, survival probability, effective sample size, Mantel-Haenszel

Kaplan Meier (KM) product-limit method has become the standard approach in estimating a survival distribution for censored data. Popular statistics are KM estimated probability of surviving beyond a specified time and median survival time. Simon and Lee's paper (1982) presents the method of obtaining CI for the probability of survival beyond a specified time. In this paper, I extend Simon and Lee's method to construct CIs for the difference of survival probabilities of two treatment groups with/without considering the strata effect. Effect sample size at that specified time is used instead of the regular sample size. The Mantel-Haenszel (MH) weighting procedure is used to adjust for strata effect. The method can be applied in oncology Phase II and III trials.

Generalized Estimating Equations for Clustered Survival Data

❖ Xiaohong Zhang, Amgen Inc., One Amgen center drive, mail stop 24 2 C, Thousand Oaks, CA 91320, *alicia70802@yahoo.com*; Kenneth Koehler, Iowa State University; Terry M. Therneau, Mayo Clinic

Key Words: clustered survival data, generalized estimating equations, Cox model, efficiency

The analysis of clustered survival data is an important statistical question. A commonly used method is to obtain the estimates of the regression parameters under the assumption of independence. A robust variance estimate is then computed which accounts for correlation. The main deficiency of the method is the loss of efficiency when the within cluster correlation is strong. We propose generalized estimating equations (GEEs) to improve the efficiency in estimating regression coefficients in the Cox model without imposing an overwhelming computational burden. Simulation was used to assess bias, variance and relative efficiency of the proposed estimators. GEEs may provide substantial gains in efficiency if within cluster correlation is sufficiently strong and the censoring rate is low. Gains in efficiency also depend on the level of homogeneity of treatment assignments within clusters.

Modeling Multicenter Survival Data: Simulation Results

❖ Daozhi Zhang, Daiichi Sankyo Pharma Development, 399 Thornall Street, Edison, NJ 08837, *daozhizhang@yahoo.com*

Key Words: Multi-Center, Survival Data, Center Effect, Clinical Trials

How to analyze multi-center survival data has been a subject of debate. The center effect is at the center of the debate. This is mainly because different approaches of handling center effect yield different statistical conclusions due to variations from center to center in patient population and number of patients. The simulation results in this presentation provide useful Knowledge on how to choose robust model from the different models including random effect model, fixed effect model, and other often used models.

A Note on Discrete-Time Survival Model with Application to Vertebral Fracture

❖ Hongwei Wang, Merck & Co., Inc., PO Box 2000, RY34 B370, Rahway, NJ 07065, *hongwei_wang@merck.com*; Weili He, Merck & Co., Inc.; Arvind K. Shah, Merck & Co., Inc.

Key Words: discrete-time survival model, proportion, power, vertebral fracture, surrogate, sample size

The status of many clinical outcomes (binary: 1/0) is assessed only at certain fixed time points with the exact time of occurrence unknown. For example in the treatment of osteoporosis, vertebral fracture is usually adjudicated through the readings of X-Ray films taken periodically. For this type of data, the proportion of patients with at least one incident at study end can be utilized as the endpoint of interest. In addition, the discrete-time survival analysis can be applied to the time to the first event. The performance of these models, in terms of power and necessary sample size, is evaluated analytically and through simulation under parameter settings that are of clinical importance to vertebral fracture trial.

247 Volatility and Risk ✪

Business and Economics Statistics Section
Tuesday, July 31, 8:30 am–10:20 am

Optimal House Tenure and Portfolio Choice with Housing as a Hedging Asset

❖ Yu Zhang, Columbia Business School, 3022 Broadway, Uris Hall 311, New York, NY 10027, *yz2112@columbia.edu*

Key Words: variable selection, portfolio choice, optimization, real estate, numerical dynamic programming

House tenure and portfolio choice are affected by several factors such as house prices, rents, labor incomes, and interest rates. Both volatilities of these factors and the comovements among them are potential factors. This paper explores the optimal life-cycle portfolio and house tenure choice with housing as a hedge against rent risks and labor income fluctuations. Forward and backward methods are employed in variable selection. Volatilities in rents and the comovement of labor income and house prices turn out to be important determinants to the household's optimization. The motive to hedge against rent risks explains the overinvestment puzzle. Constrained maximization results show that homeownership crowds out stock market participation: costly adjustments in housing calls for the liquidity provided by bonds.

Comparing Portfolio Credit Risk Methods on Diversification Effect

❖ Ming-Chin Hung, Soochow University, Dept of Business Mathematics, 56 Sec 1 KueiYang Street, Taipei, 100 Taiwan, *nhungg@bmath.scu.edu.tw*; Yi-Ping Chang, Soochow University; Huimei Liu, National Chengchi University

Key Words: Credit risk, One-factor model, Loss distribution, Value-at-Risk

One important issue in Basel II is the evaluation of portfolio credit risk. In this paper, we proposed a method to derive the exact loss distribution of a credit portfolio and calculate its value-at-risk (VaR) under one-factor model with various degrees of asset concentration. We also compare the estimated VaRs obtained from Emmer and Tasche (2005) and Basel II with the exact VaR.

Volatility in Output Growth: A Wavelet Analysis

❖ David J. Doorn, The University of Minnesota Duluth, 412 Library Drive, SBE 165, Duluth, MN 55804, *ddoorn@d.umn.edu*

Key Words: volatility, output growth, wavelet, inventories

Much recent research has documented and sought to explain declining output volatility in the U.S. since the early 1980s, with improved inventory management being one popular explanation. This study employs recently developed wavelet statistical methods to investigate the behavior of output volatility for several industrial sectors and in the aggregate, with output

viewed as the sum of sales and changes in inventories. Wavelet decomposition into scale and frequency components is used to detect change points in individual series' growth rates, while construction of wavelet covariance at different time scales reveals co-movements between series that can differ across scales. Initial results support findings of changes in volatility for sales and output taking place in the early to mid 1980s, but provide little evidence of inventory growth volatilities playing a significant role.

An Approximation Scheme for Option Pricing for Stochastic Volatility Models

❖Jungyeon Yoon, The University of North Carolina at Chapel Hill, Statistics, 214 Conner Drive, Unit 16, Chapel Hill, NC 27514, *june0821@email.unc.edu*; Chuanshu Ji, The University of North Carolina at Chapel Hill

Key Words: Option Pricing, MCMC, Stochastic Volatility, Numerical approximations

Stochastic volatility models have gradually emerged as a useful way of modeling time-varying volatility with significant potential applications, especially in finance. Stochastic volatility models alone have not proven entirely empirically successful. We use the stochastic volatility models that allow random jumps to occur in stock prices. While we keep analytical tractability that is challenged by many alternative models to Black-Scholes model by using the generalized Black-Scholes formula, we cannot avoid the computational cost that is caused by the integrals in the option pricing formula. In this paper, we propose an approximation scheme to those integrals. With MCMC algorithm, the scheme is tested and validated on simulated data. Our method is proven to be accurate and computationally much more efficient.

Empirical Analysis of Volatility Dynamics in High-Frequency Returns with a Time-Varying Component Model

❖ Ka Sing Man, Western Illinois University, IMDS Dept, College of Business and Technology, Macomb, IL 61455, *ks-man@wiu.edu*; Chunchi Wu, Singapore Management School

Key Words: high frequency returns, intraday periodicity, time-varying cyclical components, volatility persistence

We propose a new time-varying component model to analyze the intraday return volatility dynamics and the source of long-run volatility persistence. We apply the model to study IBM intraday returns. Empirical evidence indicates that the component model, consisting of a time-varying mean of absolute returns and two cosine components with time-varying amplitudes, captures very well the pronounced periodicity and persistence behaviors exhibited in the empirical autocorrelation pattern. We find that the long-run volatility persistence is driven predominantly by daily level shifts in mean absolute returns. After adjusting for these intradaily components, the filtered returns behave much like a Gaussian noise, suggesting that the model is properly specified. Furthermore, the predictive performance of using the component model for future volatility is addressed.

A Multivariate Stochastic Volatility Model and Its Inference

❖ Peng Liu, North Carolina State University, 1911 Wolf Tech Ln, Apt 304, Raleigh, NC 27603, *pliu3@ncsu.edu*; Peter Bloomfield, North Carolina State University

Key Words: multivariate stochastic volatility, nonlinear nonGaussian State-Space Model, term structure of interest rates, yield curve model, interest rate dynamics

In recent studies on financial time series, univariate and multivariate stochastic volatility (SV) models have been widely explored, but they are difficult for making inference. We consider a class of multivariate SV models. This class is based on the assumption that observed time series of length N is m-dimensional, and latent process is k-dimensional. Also assume we can observe more series and increase m, but not k. For k=1, the class reduces to the multiplicative factor model. For inference, we construct the likelihood by integrating out latent variables, using the saddlepoint approximation. Therefore, computational expense is largely reduced. By increasing m, we can enjoy the asymptotic property of the approximation. With a similar setting to that in Jacquier et al. (1994), our simulation results show that estimates are improved by increasing either m or N.

High-Dimensional Volatility Models

❖ David S. Matteson, The University of Chicago, 5455 S Ingleside Ave, Apt 3E, Chicago, IL 60615, *matteson@uchicago.edu*; Ruey S. Tsay, The University of Chicago Graduate School of Business

Key Words: Multivariate volatility, Conditional heteroscedasticity, Independent component analysis, Time-varying correlations, Leverage effect

The conditional variance, or squared volatility, of asset returns evolves over time; furthermore, financial volatilities move together over time across assets and markets. For even a handful of assets, the curse of dimensionality quickly makes estimation of most multivariate models impractical. We extend methods from Independent Component Analysis to effectively reduce the estimation problem to a set of disjoint univariate models. Our multivariate conditional heteroscedastic model allows exact or stochastic parameterizations, as well as asymmetry in the respective volatility series. Correlations evolve over time without explicit modeling, and the estimated volatility matrix is positive-definite at every time index.

248 Semiparametric Models and Biostatistics ✪

IMS, SSC, Biometrics Section

Tuesday, July 31, 8:30 am–10:20 am

Maximum Average-Power Tests

❖ Chih-Rung Chen, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, 300 Taiwan, *cchen@stat.nctu.edu.tw*

Key Words: Maximum average-power test, most powerful test, uniformly most powerful test, uniformly most powerful unbiased test

The objective of the paper is to propose and study tests that have maximum average power, averaging with respect to some specified weight function. First, some relationships between these tests, called maximum average-power (MAP) tests, and most powerful or uniformly most powerful tests are presented. Second, the existence of a maximum average-power test for any hypothesis testing problem is shown. Third, an MAP test for any hypothesis testing problem with a simple null hypothesis is constructed, including some interesting classical examples. Fourth, an MAP test for a hypothesis testing problem with a composite null hypothesis is discussed. From any one-parameter exponential family, a commonly used UMPU test is shown to be also an MAP test with respect to a rich class of weight functions. Finally, some remarks are given to conclude the paper.

Estimating the Error Distribution Function in Semiparametric Regression

❖ Ursula Mueller-Harknett, Texas A&M University, Dept of Statistics, 3143, College Station, TX 77843-3143, *uschi@stat.tamu.edu*

Key Words: local linear smoother, i.i.d. representation, efficiency

In regression models the object of primary statistical interest is the regression function. Estimators of the error distribution function are, however, also of interest, for example for tests about the regression function. There is a large literature on estimating error distribution functions, but it is nearly exclusively concerned with cases in which the regression function is parametric and can be estimated at root-n rate. We consider a partly linear regression model. Here different arguments are needed. We estimate the error distribution function by a empirical distribution function based on residuals. The residuals involve local linear smoothers. We prove a stochastic expansion for the estimator which implies a functional central limit theorem.

Inference for the Cox Regression Models with Time-Dependent Coefficients via Empirical Likelihood

❖Yanqing Sun, The University of North Carolina at Charlotte, Department of Mathematics and Statistics, Charlotte, NC 28223, *yasun@uncc.edu*; Rajeshwari Sundaram, National Institute of Child Health and Human Development; Yichuan Zhao, Georgia State University

Key Words: counting processes, empirical likelihood approach, local partial likelihood, proportional hazards model, pointwise and simultaneous confidence bands

The Cox regression models with time dependent coefficients have been studied by a number of authors recently. In this paper, we develop pointwise and simultaneous confidence bands for the time dependent regression coefficients via the empirical likelihood approach. The empirical likelihood function is formulated through the local partial likelihood for the regression coefficient functions. We show that the proposed pointwise/simultaneous confidence bands have asymptotic correct levels. Our simulation results indicate that the empirical likelihood based pointwise/simultaneous confidence bands yield better coverage probabilities than those centered at the estimators of the regression coefficient functions. The proposed method is demonstrated through a real data application.

Effects of Rounding of Conditional Frequencies on Odds Ratio

❖Juyoun Lee, The Pennsylvania State University, 3310 Plaza Drive, State College, PA 16801, *jxl982@psu.edu*; Aleksandra Slavkovic, The Pennsylvania State University

Key Words: Rounding, Log-Odds ratio, Odds ratio, Contingency Table, Conditional Frequencies

The conditional frequencies (i.e., rates) for tabular data are usually expressed as reals approximated to the true values by rounding at different decimal places. This rounding issue provides the following questions: (1) Can we say that the rounded data are still same as the unrounded data? (2) What are the effects of rounding on parameters, such as odds-ratios? These questions are highly relevant for data privacy and statistical disclosure limitation. By using tools from algebraic statistics, we investigate the difference between distributions of rounded and unrounded data. We specify the effects of the rounding on odds-ratio, and compare the bounds of cell counts and the space of the possible tables which have same conditional frequencies. We demonstrate these results with a simple example.

Use of Marginal Structural Models To Estimate the Effect of Epoetin on Survival of Dialysis Patients

❖Yi Zhang, Medical Technology and Practice Patterns Institute, 4733 Bethesda Ave Suite 510, Bethesda, MD 20814, *yz@mtppi.org*; Dennis Cotter, Medical Technology and Practice Patterns Institute; Mae Thamer, Medical Technology and Practice Patterns Institute; James Kaufman, Veterans Affairs Boston Healthcare Systems/Boston University School of Medicine; Miguel A. Hernan, Harvard School of Public Health

Key Words: Causal inference, marginal structural models, mortality

Clinical trials have shown that dialysis patients targeted to higher hematocrit levels have an increased risk of mortality. Data from United States Renal Data System were used to identify 18,353 patients who initiated dialysis and epoetin treatment in 2001. The outcome of the study is death during the 12-month follow-up period. Main exposure of interest is the average epoetin dose in the most recent three months. We conducted analyses using marginal structural models to control for time-dependent confounding arising from the treatment and hematocrit outcome. Analyses did not demonstrate a survival benefit of higher epoetin doses, consistent with RCT findings. Given the wider confidence intervals at high epoetin doses, further studies are needed to confirm the statistical significance of these findings.

Determining the Cutoff Points for a Continuous Biomarker When Its Interaction with Study Treatment Is of Interest with Time-to-Event Outcomes

❖Keyue Ding, Queen's University, 10 Stuart, Kingston, ON K7L 3N6 Canada, *kding@ctg.queensu.ca*

Key Words: Clinical trials, Martingale residuals, Predictive factor, Biomarker, Cutoff points

Increasing knowledge in cancer biology leads to more accurate diagnosis, prognosis assessment and development of targeted therapy. To maximize patients benefit in apply those achievements, an important issue is to identify predictive marker in treatment selection, i.e., to treat patients according to the level of the predictive marker. In this paper, we propose a graphic approach to determine the cutoff points for a continuous marker, according to the treatment effects in intervals of the marker value. Most important part is determined at which range the new treatment is no better than standard treatment/placebo. Relative issues in determining the optimal cutoff points are discussed. Finally, we apply the method in categorized biomarkers for treatment selection in two NCIC CTG lung cancer trials.

Robust Methods for Estimation of State Occupancy Probabilities with Interval-Censored Multistate Data

❖David Tolusso, University of Waterloo, Statistics and Actuarial Science, 200 University Ave W, Waterloo, ON N2J 4H2 Canada, *dctolusso@math.uwaterloo.ca*; Richard J. Cook, University of Waterloo

Chronic disease processes are often conveniently modeled using multistate models. In many settings however, patients are only observed at periodic assessment times in which case transition times are interval censored. In this talk we consider robust methods for estimating state occupancy probabilities based on interval-censored multistate data. Methods are described based on marginal failure time models and Markov models fitted via local likelihood. Data from a study of patients with psoriatic arthritis are used for illustration.

249 Nonparametric Methods for Machine Learning and High-Dimensional Data ✪

Section on Nonparametric Statistics

Tuesday, July 31, 8:30 am–10:20 am

A Method for Dimension Reduction in Classification with Functional Data

❖Santiago Velilla, Universidad Carlos III de Madrid, Calle Madrid 126, Getafe Madrid, 28903 Spain, *santiago.velilla@uc3m.es*

Key Words: basis expansions, canonical correlation analysis, functional dimension-reduction subspaces

A method for dimension reduction in discriminant analysis with functional data is presented. The basic idea is to generalize the concept of dimension-reduction subspace in the finite-dimensional case to the context in which the observations arise from a stochastic process. A well-known data example is analyzed. Comparisons with other approaches in this area are also given.

Exact Bagging of k-Nearest-Neighbor Learners

❖Brian Steele, University of Montana, Dept of Mathematical Sciences, Missoula, MT 59812, *steele@mso.umt.edu*

Key Words: bagging, k-nearest neighbor, classification, prediction, bootstrap, statistical learning

Bootstrap aggregation, or bagging, is a method of reducing the prediction error of a statistical learner. In the context of the prediction problem, the goal of bagging is to construct a new learner which is the expectation of the original over the empirical or sample distribution function. In nearly all cases, the expectation cannot be computed analytically, and bootstrap sampling is used to produce an approximation. The k-nearest neighbor learners are exceptions to this generalization, and exact bagging is very easy. In addition to computational savings there are interesting opportunities to study the bagging properties of k-nearest neighbor learners and to develop new exact and nearly exact bootstrap k-nearest neighbor learners.

Consistency of Lasso Selection in High-Dimensional Approximating Models

Florentina Bunea, Florida State University; ❖Shuva Gupta, Florida State University, Tallahassee, FL 32306-4330, *shuva@stat.fsu.edu*

Key Words: LASSO, Oracle estimator, Penalized Regression, asymptotic normality, approximating models, \$l_{1}\$ penalty

Here we study the consistency property of a subset of covariates selected using a \$\ell_{1}\$ penalty in high dimensional approximating models. Invoking the result by Bunea et al. (2006), which states that the covariates selected by penalized least squares estimator satisfies sparsity oracle inequalities, we prove that the covariates selected using an \$\ell_{1}\$ penalty converges to an oracle target which acts as the ``true” underlying set of covariates in nonparametric regression setting with random design. The results are valid even when the dimension of the model is much larger than the sample size. This result can be applied to high dimensional linear regression or when the design matrix is not invertible.

Statistical Approaches to Mining Multivariate Datastreams

❖Eric Vance, Duke University, 114 Old Chemistry Building, Box 90251, Durham, NC 27708, *ervance@stat.duke.edu*

Key Words: data streams, depth-based partitioning , multivariate histograms, Mahalanobis distance

Datastreams are fast moving, rapidly accumulating sequences of data that are a predominant source of information today. Monitoring these datastreams for anomalous behavior faces special challenges—rapid rate of accumulation; massive size and complexity; and brief access to the raw data. An effective strategy to address these issues is to first partition the attribute space into meaningful classes, and then use the class-summaries to analyze the streams further. In the context of two real-life applications we discuss a variety of data partitioning schemes and compare their performance in answering a fundamental question in the analysis of datastreams: Has the statistical behavior of the datastream changed? If so, in what way?

Modeling Computer Experiments with Multiple Responses

❖Zhe Zhang, The Pennsylvania State University, 4048E Bannock Burn place, Charlotte, NC 28211, *zxz118@stat.psu.edu*; Runze Li, The Pennsylvania State University; Agus Sudjianto, Bank of America

Key Words: computer experiment, kriging model, functional ANOVA

This paper is concerned with modeling computer experiments with multiple responses. It has been common that one may collect multiple responses from a physical or computer experiment. However, to our best knowledge, there is little work to model computer experiment with multiple responses. In this paper, we propose a modeling procedure for such computer experiment by using a multivariate kriging model, a natural extension of the ordinary kriging model. We further extend functional ANOVA of single response to multiple response and apply it for analyzing the effect of each design variable. The proposed methodology is demonstrated by an analysis of data collected from a case study concerned with the design of the engine structure to minimize the radiated noise.

Nonparametric Tests and Finite Approximations for Balanced and Unbalanced Multivariate Data, with Applications

❖Arne Bathke, University of Kentucky, Department of Statistics, 875 Patterson Office Tower, Lexington, KY 40506-0027, *arne@ms.uky.edu*; Solomon W. Harrar, University of Montana

Key Words: Multivariate Tests, Nonparametric Methods, Unbalanced Design

We consider multivariate data from different experiments (e.g., plant disease response data). To analyze this type of data, we propose different nonparametric tests for multivariate observations in balanced and unbalanced one-way layouts and present the corresponding asymptotic results. In addition, we investigate several small-sample approximations and compare their performance for different simulated situations with continuous and discrete observations, including high-dimensional data.

Methodological Aspect of Interpretation of Principal Components

❖Lev Sverdlov, 200 Baxter Blvd 3, Portland, ME 04101, *lev.sverdlov@gmail.com*

Key Words: Factor analysis, Principal components, Interpretation, Content, Meaning

This report describes a strategy facilitating translation of statistical results of principle components analysis (PCA) into a language of content and meaning (i.e., “surmising” the content and meaning of the statistically identified principal component). The strategy includes 1) presenting variables in a comparative way (“younger” vs. “older” instead of “age”); 2) the development of two contrasting (positive and negative) but logically equivalent presentations of each retained component; 3) sorting absolute values of loadings in a descending order, and 4) contrasting positive and negative presentations of the factor. The graphical presentation can be helpful for the interpretation. The strategy is illustrated by the results of studies (Sverdlov et al., 1993, 2004). Psychological and philosophical aspects of the interpretation of PC are discussed.

250 Computer Experiments and Network Traffic Models ● ✪

Section on Physical and Engineering Sciences

Tuesday, July 31, 8:30 am–10:20 am

Simultaneous Calibration and Tuning of Computer Experiments

❖Gang Han, The Ohio State University, 2517 Burlawn Court, Columbus, OH 43235, *han.191@osu.edu*; Thomas Santner, The Ohio State University

Key Words: Hierarchical Bayesian Model, product power exponential correlation, kriging, Metropolis Hastings algorithm, optimization, root mean squared prediction error

Calibrating complex computer codes to field data and setting tuning parameters in computer outputs are both problems of considerable interests to researchers. We propose a methodology that optimizes the tuning parameters, estimates the distribution of calibration variables, and predicts the untried physical experiment responses. A Bayesian model is constructed based on conditional Gaussian stochastic processes and diffuse prior distributions. The model is implemented by Markov Chain Monte Carlo methodology. The program is illustrated with an example and applied in a biomechanics engineering problem.

Estimating Percentiles in Computer Experiments

❖Soma Roy, The Ohio State University, 1958 Neil Avenue, 404 Cockins Hall, Columbus, OH 43210, *somaroy@stat.ohio-state.edu*; William Notz, The Ohio State University

Key Words: Expected improvement criterion, Sequential estimation, Efficiency within a budget

This paper discusses percentile estimation in a computer experiments setting. We consider the case with multiple inputs and a single output. We assume that the inputs have some known distribution and wish to estimate the p-th percentile of the induced distribution of the output. We propose a sequential procedure based on an expected improvement criterion that will perform efficiently within a fixed budget. The algorithm is a two-step procedure. At stage n, training data are used to first estimate the p-th percentile of the output, say Yp,n, and then estimate the input variable settings that would produce Yp,n. The expected improvement criterion is used to select the next point to add to the training data to improve our estimate of the p-th percentile. We observe the computer code at this point, and iterate to stage n+1. Our procedure is based on previous work by Ranjan, P. et al. (2006).

Blind Kriging

❖Roshan Vengazhiyil, Georgia Institute of Technology, 755 Ferst Dr NW, Atlanta, GA 30332, *roshan@gatech.edu*; Ying Hung, Georgia Institute of Technology; Agus Sudjianto, Bank of America

Key Words: Computer experiments, Finite element models, Kriging, Metamodels, Product design, Variable selection

Kriging is a useful method for developing metamodels for product design optimization. The most popular kriging method known as ordinary kriging uses a constant mean in the model. In this article, a modified kriging method is proposed, which has an unknown mean model. Therefore it is called blind kriging. The unknown mean model is identified from experimental data using some variable selection techniques. Many examples are presented, which show remarkable improvement in prediction using blind kriging over ordinary kriging. The blind kriging predictor is easier to interpret and also seems to be more robust to the misspecification in the correlation parameters.

Optimal Design of Network Traffic Measurement Scheme

❖Harsh Singhal, University of Michigan, 1929 Plymouth Road, Apt 4038, Ann Arbor, MI 48105, *singhal@umich.edu*; George Michailidis, The University of Michigan

Key Words: Optimal Design, Computer Network, Traffic Volume Measurement

Computer network traffic can be sampled at routers to determine the origin and destination of different packets. This is in turn used to estimate the volume of traffic between all origin-destination pairs. Use of higher sampling rate leads to better estimates but uses more resources. In this work we formulate the problem of determining the network-wide sampling rates as an optimal design problem. This problem is solved using semi-definite programming and simulations show that this design performs better than conventional sampling schemes.

Sample-Based Estimation of Internet Traffic Flow Characteristics

❖Lili Yang, The University of Michigan, 439 West Hall 1085 South University, Ann Arbor, MI 48109-1107, *yanglili@umich.edu*; George Michailidis, The University of Michigan

Key Words: EM Algorithm, Network, nonparametric, flow, internet, traffic

Understanding the characteristics of traffic flows is crucial for allocating the necessary resources (bandwidth) to accommodate users demand. In this paper, the problem of nonparametric estimation of network flow characteristics based on sampled flow data from single-stage Bernoulli sampling and two-stage sampling will be addressed. An adaptive Expectation-maximization (EM) algorithm is used for the flow length distribution, which in addition provides an estimate for the number of active flows. The estimation of the flow sizes (in bytes) is accomplished through a regression model. The proposed approaches are illustrated and compared on a number of synthetic and real datasets.

Propagating SLAs Statistics in the IT Service Provider Environment

❖Larisa Shwartz, IBM T.J. Watson Research Center, 19 Skyline Dr, Hawthorne, NY 10532, *lshwart@us.ibm.com*; Genady Grabarnik, IBM T.J. Watson Research Center

Key Words: enterprise analytics, business process modeling, workflow, automation

IT Service Providers are facing an increasingly intense competitive landscape and growing government and industry requirements such as standards and best practices. In their quest to maximize customer satisfaction, Service Providers seek to employ business intelligent solutions which provide deep analysis and orchestration of business processes and capabilities for optimizing the level of service and cost. In this paper we give a formal description of the SLA on different Service Provider/ Customer/ Supplier (SPCS) models, provide a formula for the calculation of different SLA related statistics over a formal business processes, show how different SPCS models impact resulting statistics and SLAs, and outline way of finding majorating SLAs for a Service Provider for various models.

Estimating Traffic Volume over a Large Transportation Network from Noisy Data

❖Jaimyoung Kwon, California State University, East Bay, Dept of Statistics, Hayward, CA 94542, *jaimyoung.kwon@csueastbay.edu*; Karl Petty, Berkeley Transportation Systems, Inc.

Key Words: network, graph theory, flow conservation, weighted least squares, space-time data

Accurate measurement of traffic volumes throughout a transportation network is critical for efficient traffic management and as well as for simulation studies. Obtaining a network-wide volume estimate is a challenging task since: a) transportation networks are large and complex; b) measurements from many sensors, each with different noise level, need to be fused together; and c) the final estimate needs to globally satisfy the flow conservation constraint. We develop a graph-theoretic, weighted least squares regression (WLSR) approach for the problem. The approach uses graph theory to construct the design matrix for WLSR, which represents the flow conservation equation. Also, measurement errors are reflected a priori as the weight matrix for WLSR. The performance of the algorithm is studied via simulation and application to data from real freeway networks in Greece and California.

251 Machine Learning ●

Section on Statistical Computing, Section on Statistical Graphics, Section on Physical and Engineering Sciences, Section on Quality and Productivity

Tuesday, July 31, 8:30 am–10:20 am

Semisupervised Learning from Dissimilarity Data

❖Michael Trosset, Indiana University, PO Box 6424, Bloomington, IN 47407, *mtrosset@indiana.edu*; Carey E. Priebe, Johns Hopkins University

Key Words: embedding, multidimensional scaling, principal components, classification, discriminant coordinates

One approach to classifying objects for which only pairwise dissimilarities are available is to (1) embed the objects in Euclidean space (e.g., by classical multidimensional scaling), then (2) apply a conventional classification procedure (e.g., linear discriminant analysis). This two-stage approach is semisupervised: although only labeled objects can be used to construct the classifier, additional unlabeled objects can be used to facilitate construction of the representation space. We explore the extent to which a semisupervised approach improves on a fully supervised approach.

Algorithms for Support Vector Machines

❖Denise Reeves, George Mason University, 9984 Hemlock Woods Lane, Burke, VA 22015, *dreeves@gmu.edu*

Key Words: ill-posed inverse problems, bias and variance dilemma, optimal coordinate system

Estimation of sufficient statistics and parameters are ill-posed inverse problems characterized by partially specified system models. Such problems require mathematical constructs that rule out physically implausible solutions with simple, eloquent representations. Classical estimation techniques fail to mitigate the bias and variance dilemma. The right structure or bias is often an optimal coordinate system. The likelihood ratio is generally reduced to a sufficient statistic. We consider the optimal curve and coordinate system being estimated by the support vector algorithm and precisely clarify the nature of the bias for common covariance matrices.

Random Forests for Feature Selection: To Be Handled with Caution

❖Carolin Strobl, LMU Munich, Ludwigstr 33, Dept. of Statistics, Munich, 80539 Germany, *carolin.strobl@stat.uni-muenchen.de*

Key Words: feature selection, random forest, bagging, subagging, bootstrap, bias

Variable importance measures for random forests are receiving increasing attention as a screening tool in high dimensional classification tasks, e.g. in statistical genomics. However, we show and illustrate in simulation studies that suboptimal predictor variables may be artificially preferred by these measures when predictor variables vary in their scale level or their number of categories. The two statistical mechanisms underlying this deficiency are biased variable selection in the individual classification trees on one hand, and effects induced by bootstrap sampling with replacement on the other hand. An alternative implementation of random forests and bagging, providing unbiased variable selection, is presented. When this method is applied with subsampling without replacement, the resulting variable importance measure can be used reliably as a screening tool in any data situation.

Feature Selection for Large Data

❖Peng Liu, Case Western Reserve University, 26241 Lake Shore Blvd, Apt 1951, Euclid, OH 44132, *pxl59@case.edu*; Jiayang Sun, Case Western Reserve University

Key Words: data mining, large data, feature selection, mixture distribution, partial EM, intrusion detection

A typical challenge in data mining is that data can be too large to be loaded into a computer program once and for all for an analysis, or data come sequentially in streams, so it is necessary to work on pieces of the data and then combine the information from different pieces to obtain the whole picture. In this talk we describe our recent research in developing techniques for feature selection and mixture estimation for large data. Their performance is evaluated by asymptotic analysis and simulation, and compared with standard algorithms. The application of our proposed methods in intrusion detection is demonstrated on the KDD Cup 1999 dataset. (Part of the talk is based on joint work with J. Chen and Z. Zhang.)

Asymptotic Mean Squared Prediction Error of L2Boosting Estimator Under Mis-specified Models

❖Tzu-Chang Cheng, University of Illinois at Urbana-Champaign, Department of Economics, No 1306 North Lincoln Avenue, Apt 318, Urbana, IL 61801, *r92323047@ntu.edu.tw*; Ching-Kang Ing, Academia Sinica

Key Words: Asymptotic Mean Squared Prediction Error, L2Boosting, Weak greedy algorithm, Weak greedy orthogonal algorithm

A rigorous analysis of the L2Boosting predictor is given in a high-dimensional linear regression model, in which the number of regressors is much larger than the number of observations. Instead of boosting the predictor using the weak greedy algorithm (WGA) in Buhlmann (2006), this paper focuses on the weak orthogonal greedy algorithm (WOGA) because its statistical properties are more traceable. We explore the variable sequences determined by WOGA, and thereby give an asymptotic expression for the mean-squared prediction error (MSPE) of the corresponding Boosting predictor. This expression shows an algebraic tradeoff between the bias and variance terms. Finally, our simulation results suggest that the Boosting predictor based on WOGA outperforms the one based on WGA in terms of MSPE, offering an interesting direction for future research.

On Efficient Supervised Learning of Multivariate t Mixture Models with Missing Information

❖Tsung-I Lin, National Chung Hsing University, Department of Applied Mathematics, Taichung, 402 Taiwan, *tilin@amath.nchu.edu.tw*; Hsiu-J Ho, National Chung Hsing University; Pao-S Shen, Tunghai University

Key Words: Classifier, Learning with missing information, Missing values, Multivariate t mixture model, PX-EM algorithm, Outlying observations

A finite mixture model using the multivariate t distribution has been well recognized as a robust extension of Gaussian mixtures. This paper presents an efficient PX-EM algorithm for supervised learning of multivariate t mixture models in the presence of missing values. To simplify the development of new theoretic results and facilitate the implementation of the PX-EM algorithm, two auxiliary indicator matrices are incorporated into the model and shown to be effective. The proposed methodology is a flexible mixture analyzer that allows practitioners to handle real-world multivariate data sets with complex missing patterns in more efficient manners. The performance of computational aspects is investigated through a simulation study and the procedure is also applied to two real data sets with varying proportions of synthetic missing values.

Conditional Confidence Intervals for Classification Error Rate

❖ Hie-Choon Chung, Gwangju University, Dept of e-Business, 592-1 Jinwoldong Namgu, Gwangju, 503-703 South Korea, *hcc@gwangju.ac.kr*; Chien-Pai Han, University of Texas at Arlington

Key Words: Conditional Jackknife Confidence Interval, Conditional Bootstrap Confidence Interval, Conditional Error Rate, Monte Carlo Study

An observation is to be classified into one of several multivariate normal populations with equal covariance matrix. When the parameters are unknown, independent training samples are taken from the populations. We consider the construction of confidence interval for the conditional error rate. The cases of two populations and three populations are studied in detail. The conditional jackknife confidence interval and the conditional bootstrap confidence intervals of the conditional error rate are obtained and compared.

252 Variance Estimation: Advanced Applications ●

Section on Survey Research Methods

Tuesday, July 31, 8:30 am–10:20 am

Quasi–Monte Carlo Methods for Variance Estimation

❖ Stanislav Kolenikov, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65211, *kolenikovs@missouri.edu*

Key Words: design-based inference, balanced repeated replication, bootstrap, quasi Monte Carlo, Halton sequence

In this talk, I shall consider the application of quasi-Monte Carlo tools, such as Halton sequences, to create approximately balanced replication designs. The QMC methods are designed to simulate points in multidimensional unit cubes with coverage asymptotically more uniform that that of the standard Monte Carlo. I will adopt those methods to generate replication designs that occupy an intermediate position between highly balanced designs such as BRR, and random replication designs such as the bootstrap. Thus the resulting designs can be viewed as a viable alternative to the extensions of BRR based on mixed orthogonal arrays, or to the balanced bootstrap.

Estimating the Distribution of Distances for a Large-Scale Complex Survey

❖ Jianqiang Wang, Iowa State University, CSSM, 204 Snedecor, Ames, IA 50011, *jqwang@iastate.edu*; Jean Opsomer, Iowa State University

Key Words: distance metric, kernel estimation, empirical distribution, estimating equations, generalized median

Many finite populations targeted by sample surveys consist of homogenous subpopulations with respect to the variables being collected. We propose a sample-based estimator for the subpopulation distribution functions of the distances between the elements and the subpopulation centers, and explore several definitions of distance in the data and different ways to define the subpopulation centers. We describe the theoretical properties of the estimator, and propose a variance estimator. We discuss a procedure to identify outliers in large-scale surveys that takes advantage of the estimated subpopulation distribution functions, and apply it to data from the National Resources Inventory.

Variance Estimation for Calibration with Estimated Control Totals

❖ Jill Dever, University of Maryland, Joint Program in Survey Methdology, 1218 Lefrak Hall, College Park, MD 20742, *jdever@survey.umd.edu*; Richard Valliant, University of Maryland

Key Words: estimated-control calibration, survey-estimated control totals, variance estimation, simulation

Calibration estimators use auxiliary information to improve the efficiency of survey estimates. The control totals, to which sample weights are calibrated, are assumed to be population values; often, however, the controls are estimated from other surveys. Many researchers apply traditional calibration variance estimators to cases where the control totals are estimated, thus assuming that any additional sampling variance associated with these controls is negligible. The goal of our research is to develop and evaluate variance estimators for two-stage designs with estimated-control (EC) calibration using design-unbiased controls. We compare linearization and replication variance estimators via simulation for point estimators of totals and means. Illustrations are given of the effects on variances of different levels of precision in the estimated controls.

Variance Estimation After Nearest-Neighbor Imputation with Application to 2000 Census Long Form Sample

❖ Jae-kwang Kim, Yonsei University, Department of Applied Statistics, Seoul, 120-749 South Korea, *kimj@yonsei.ac.kr*; Wayne Fuller, Iowa State University

Key Words: poverty, income, item nonresponse

The Census 2000 long form is used to construct estimates of income and poverty at the state, county, and school district levels. Variance estimation for the estimators must account for (1) uncertainty due to imputation, and (2) raking to census population controls. The Kim and Fuller (2004) method for variance estimation under nearest neighbor imputation is adapted to this problem. Issues related to variance estimation for domains are discussed.

On Variance Estimation Under Auxiliary Value Imputation in Sample Surveys

❖ David Haziza, University of Montréal, CP 6128 succ Centreville, Montréal, QC H3C3J7 Canada, *david.haziza@umontreal.ca*; Jean-François Beaumont, Statistics Canada; Cynthia Bocci, Statistics Canada

Key Words: Cold-deck imputation;, Imputation model, Nonresponse model, Response probability, Robust variance estimator, Self-efficiency

In this paper, we study the problem of variance estimation for a domain total when auxiliary value imputation, sometimes called cold-deck or substitution imputation, has been used to fill in missing data. We consider two approaches to inference, which lead to different variance estimators. In the first approach, the validity of an imputation model is required. Our proposed variance estimator is nevertheless robust to misspecification of the second moment of the model. Under this approach, we show the somewhat counter-intuitive result that the total variance of the imputed estimator can be smaller than the sampling variance of the complete-data estimator. We also show that the naÔve variance estimator (i.e. the variance estimator obtained by treating the imputed values as observed values) is asymptotically a valid estimator of the total variance when the sampling fraction is negligible.

The Effect of Stochastic Population Controls on the Precision of Radio Listening Estimates

❖Richard Griffiths, Arbitron Inc., 9705 Patuxent Woods Drive, Columbia, MD 21044, *richard.griffiths@arbitron.com*

Key Words: Weighting, Raking, Precision, Stochastic population controls, Radio

Arbitron uses a raking procedure to weight the sample from its quarterly surveys to external population controls. Some of these population controls are based on a survey with relatively small sample sizes. They thus contribute to the sampling error found in the final radio listening estimates. In this paper, we examine the results of a detailed empirical investigation into the effect of the stochastic population controls on the precision of the estimators.

253 Improved Methods for Survey Contact and Incentives ●

Section on Survey Research Methods

Tuesday, July 31, 8:30 am–10:20 am

Using Linear Programming Techniques for Balancing Automatic Case Assignments in Field Surveys

❖Edward English, NORC at the University of Chicago, 55 E Monroe St, Suite 4800, Chicago, IL 60603, *english-ned@norc.org*; Steven Pedlow, NORC at the University of Chicago

Key Words: Area Probability Sample, Interviewer allocation

National field surveys attempt to hire field interviewers who reside proximate to selected cases. Traditionally, case assignment was a very labor-intensive activity, with staff manually comparing national maps to determine where cases are in comparison with interviewers. NORC has been using an automatic distance-based case assignment procedure that generally assigns cases to the closest interviewer. However, many interviewers have been allocated unreasonably high work loads, while others had been assigned none. The current research uses linear programming to create constraints during allocation in order to lessen the degree of manual intervention and thus balance loads automatically. Our method realized an increase in operational efficiency without a significant loss in accuracy, and thus shows promise as a method for integrating GIS, programming, and survey research.

First Impression: An Advance Contact Experiment to Locate and Engage Potential Respondents

❖Michael E. White, U.S. Census Bureau, 4600 Silver Hill Rd. HQ 7H028D, Washington, DC 20233-8732, *Michael.e.white@census.gov*; Megan Henly, U.S. Census Bureau; Amy B. Herron, U.S. Census Bureau; Andrew Zukerberg, U.S. Census Bureau

Key Words: nonresponse, locating mobile populations, advance contact methods, NSRCG

Locating sampled individuals is an important first step for surveys with a young, mobile target population like the National Survey of Recent College Graduates (NSRCG). In previous cycles, locating was done by mailing an address verification form (TAVF) to the graduate's last known address. After locating graduates, another important step is to engage them by describing the importance of the survey. The 2006 NSRCG compared the effect of mailing a TAVF to that of a colorful brochure requesting the same information but also highlighting results from previous surveys. Three brochures were developed: one targeted graduates of degree fields with high nonresponse; a second catered to previously low responding racial minorities; a third provided information relevant to all graduates. This paper will evaluate the impact of the four mailings on obtaining updated addresses and overall response.

The Impact of Incentives on Survey Participation and Reports of Intimate Partner and Sexual Violence

❖Lisa Carley-Baxter, RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 22709-1294, *lcbaxter@rti.org*; Michele Lynberg, Centers for Disease Control and Prevention; Susan Twiddy, RTI International

Key Words: incentives, experiment, violence, RDD, respondent reactions

An incentive experiment is being used in a Pilot Study for the National Intimate Partner and Sexual Violence Survey (NISVS). The NISVS Pilot is a nationwide random-digit-dial survey that collected sensitive information on the perpetration and victimization of five types of violence from adults 18 years of age and older. Sampled telephone numbers were preassigned to receive either a $10 or $20 incentive payment upon completion of the interview. In this paper, we compare response rates, respondent demographics, and substantive differences on particular indicators by incentive group. Due to the sensitive nature of the survey questions, the NISVS pilot also includes a series of questions about the respondent's reactions to the survey. We analyze the potential impact of differing incentives on the respondent's reactions to questions about intimate partner violence and sexual violence.

Money Can Buy Me Love: Experiments To Increase Response Rates Using Monetary Incentives

❖Andrew Zukerberg, U.S. Census Bureau, 4600 Silver Hill Rd., Washington, DC 20233, *andrew.l.zukerberg@census.gov*; Megan Henly, U.S. Census Bureau; David Hall, U.S. Census Bureau

Key Words: incentives, surveys

This paper focuses on three experiments designed to increase survey response through the use of monetary incentives. These experiments were conducted on two surveys of the college-educated population. In the first experiment, we targeted "potential refusals" and "final refusals." The next experiment targeted low responding groups. The final experiment focused on noncontact sample cases. In all experiments, cases were randomly assigned to either a control or treatment group. The control group received a letter encouraging survey participation and the treatment group received an inactive $20 gift card with the letter. The treatment letter indicated that the gift card would be activated upon completion of the survey. In this paper, we examine the effect the incentive had on response in each experiment and make recommendations on the use of incentives in future survey years.

254 Teaching Statistics in the Health Sciences

Section on Teaching Statistics in the Health Sciences, Section on Statistical Education

Tuesday, July 31, 8:30 am–10:20 am

The Job Outlook for Biostatistics Graduates

❖Stephen W. Looney, Medical College of Georgia, Dept of Biostatistics, 1120 15th Street AE3020, Augusta, GA 30912-4900, *slooney@mcg.edu*; Joseph L. Hagan, Louisiana State University Health Sciences Center

Key Words: employment opportunities, Amstat News, Masters students, Ph.D. students

In this presentation, we consider the problem of estimating the annual number of masters and doctoral biostatistics graduates in the United States. We also consider estimation of the number of job opportunities for these graduates. We describe and compare several methodologies used to produce these estimates. We discuss the inherent difficulties in estimating the number of entry-level jobs available. We also compare our methodology to that of a similar study conducted in 1994. We conclude that the job outlook is good for biostatistics graduates, especially for those obtaining a PhD.

Developing a Biostatistical Collaboration Course in a Health Research Methodology Program

❖Lehana Thabane, McMaster University, Centre for Evaluation of Medicines, 105 Main Street Level P1, Hamilton, ON L8N 1G6 Canada, *ThabanL@mcmaster.ca*; Stephen Walter, McMaster University; Steven E. Hanna, McMaster University; Charlie Goldsmith, McMaster University; Eleanor Pullenayegum, McMaster University

Key Words: Biostatistics, Health Research, Collaboration, Mentor, Internship, Graduate course

Effective statistical collaboration in a multidisciplinary health research environment requires skills not taught in theoretical statistics courses. Graduates often learn such skills through trial-and-error. In this presentation, we discuss the development of a biostatistical collaboration course aimed at graduate level students in a Health Research Methodology PhD program with Specialization in Biostatistics. The objectives of the course include to promote enthusiasm for and commitment to excellence in statistical collaboration in health research, to develop skills in applied statistics, and to enhance collaboration and communication among statisticians and nonstatistician researchers. The course uses an innovative combination of lectures, tutorial discussions with guest lecturers, videotaped practice sessions, and internship with an experienced biostatistician mentor.

The Effect of Effect Size on Teaching, Collaboration, and Publishing

❖Linda S. Hynan, The University of Texas Southwestern Medical Center at Dallas, Division of Biostatistics, 5323 Harry Hines Blvd, Dallas, TX 75390, *linda.hynan@utsouthwestern.edu*; Alan Elliott, The University of Texas Southwestern Medical Center at Dallas

Key Words: Effect Size, p-values, hypothesis testing

Statistical significance testing has been an essential tool of researchers for decades. Hypothesis testing provides a way to remove personal bias from the decision-making process. However, overconfidence in decisions based on p-values less than 0.05 has undoubtedly led to many erroneous conclusions. The rampant use of tests on the same data set guarantees that incorrect decisions are made. Furthermore, a strict interpretation of significance testing without regard to scientific importance leads to missed opportunities for scientific advancement and may result, for example, in the implementation of an ineffective medical procedure or the abandonment of a line of research that could have proven valuable. This paper reviews the pros and cons of reliance on p-values and how APA guidelines that recommend reporting effect sizes could change how results are presented in scientific journals.

Use of Interesting Examples from Medicine and Biology

❖Harry Norton, Carolinas Medical Center, 5500 Lansing Drive, Charlotte, NC 28270, *jnorton@carolinas.org*

Key Words: Teaching, Statistical Education, Introductory Biostatistics, Examples from Medicine

Keeping undergraduate biology students and medical residents interested in statistics, when most of the students are taking the class as a requirement can be challenging. I have attempted to accomplish this goal by using real-life and engaging examples from medicine and biology. I will present a number of examples that I use in my course instruction. These will include published medical papers containing critical mistakes, fallacies in numerical reasoning, the Will Roger's Phenomenon, and population estimation using the capture/recapture method.

WMWodds: A Simple and Direct Effect-Size Parameter for the Wilcoxon-Mann-Whitney Two-Sample Test

❖Ralph G. O'Brien, The Cleveland Clinic, Quantitative Health Sciences WB4, Cleveland, OH 44195, *OBrienRalph@gmail.com*; John Castelloe, SAS Institute Inc.

Key Words: Wilcoxon two-sample test, Mann-Whitney test, power, effect size

The venerable Wilcoxon-Mann-Whitney (WMW) two-sample test has outstanding utility, but is insufficiently understood by many professional statisticians and thus is too often incorrectly taught, interpreted, and reported. Much of this confusion stems from pretending that the two distributions have identical shapes and spread, but possibly different locations, a model that rarely resembles reality. A simple odds parameter, WMWodds, allows us to express the statistical hypotheses correctly and to obtain point estimates and confidence intervals that are directly related to the WMW test, yet are not dependent on the location-shift model. This methodology also yields a new approximation for the WMW p-value and a new noncentrality parameter for performing power/sample-size analyses.

Calibrated Peer Review (CPR) as a Tool To Teach Interpretation of Linear Regression Coefficients: Development of the CPR Module and Examination of Students' Results

❖Felicity B. Enders, Mayo Clinic, 200 First Street SW, Ha 706, Rochester, MN 55905, *enders.felicity@mayo.edu*; Sarah Stoner, Mayo Clinic; Verna Hoverman, Mayo Clinic

Key Words: Statistics education, Calibrated peer review, Regression

Linear regression can be an extremely challenging topic. In this class of 27 masters of clinical research students, prime importance was placed on learning to correctly interpret regression parameters. In order to facilitate learning, we developed a calibrated peer review (CPR) module (http://cpr.molsci.ucla.edu) to help students interpret coefficients for a single binary predictor, a single continuous predictor, or a set of dummy variables. Students' interpretation scores improved after completing the module ($p=0.06$). Evaluation data show that while some students did not enjoy using the CPR system, they did feel it helped them to learn ($p=0.04$). In particular, students reported that grading their own initial assignment was a helpful learning exercise ($p=0.002$). Calibrated Peer Review can be a useful learning tool for text-based statistical exercises.

A SAS Macro Implementation of a Multiple Comparison Post Hoc Test

❖Alan Elliott, The University of Texas Southwestern Medical Center at Dallas, Division of Biostatistics, 5323 Harry Hines Blvd, Dallas, TX 75390, *alan.elliott@utsouthwestern.edu*; Linda S. Hynan, The University of Texas Southwestern Medical Center at Dallas

Key Words: Kruskal-Wallis, nonparametric, SAS, post hoc, multiple comparisons, Dunn

The Kruskal-Wallis non-parametric analysis of variance is often used instead of a standard ANOVA when samples come from a suspected non-normal population or when variances are non-heterogeneous. This omnibus procedure tests for differences between groups, but provides no specific post hoc pairwise comparisons. This paper provides a SAS macro implementation of a Tukey-type multiple comparison test (Dunn's Test) based on Kruskal-Wallis results from the SAS NPAR1WAY procedure. The implementation can be utilized for up to 10 groups at the 0.05 alpha significance level. In addition, the effectiveness of the procedure to meet a selected type I significance level is evaluated using Monte Carlo simulations.

255 Estimating Social Statistics: Hurricanes, Murders, and Other Applications ● ✪

Social Statistics Section

Tuesday, July 31, 8:30 am–10:20 am

Assessing Katrina's Impact: Report of a Study on the Mississippi Gulf Coast

❖David A. Swanson, University of Mississippi, Sociology Anthropology Dept, University, MS 38677-1848, *dswanson@olemiss.edu*; Mark Van Boening, University of Mississippi; Richard Forgette, University of Mississippi

Key Words: Katrina, sample, census

This study reports the results of a two-pronged study funded by the National Science Foundation to examine the effects of Hurricane Katrina. The project is focused on the area of the Mississippi Gulf Coast immediately to the west and east of Bay St. Louis. In terms of our study goals: (1) we report the results of testing hypotheses about the effect of social networks on the well-being of people in the study area who were impacted by Hurricane Katrina; and (2) we provide a systematic count of the stock of housing in the study area and compare it with pre-Katrina information to obtain a picture of damage consistent with census concepts and definitions.

Factors That Facilitated and Inhibited Job-Holding Among Female AFDC/TANF Recipients in 1996

❖Denton Vaughan, Consultant, 11237 Handlebar Road, Reston, VA 20191, *drvaughan@comcast.net*

Key Words: Welfare reform, employment, impediments, facilitating factors, multiple

This study employs the Survey of Income and Participation (SIPP) to assess the probability that AFDC/TANF recipients worked the initial year of welfare reform. A set of rather common place attributes were subjected to bi-variate statistical tests to assess their impact on jobholding of primary recipients. This simple descriptive approach shows that those who experienced one or more impediments accounted for more than four-fifths the national caseload. Of the more than one-quarter of primary recipients reporting three or more impediments, only 21 percent held a job in 1996. On the other hand, two-fifths of AFDC/TANF recipients reporting three or more factors that facilitated work held a job at some point during they year. Thus the AFDC/TANF caseload was very differentiated in terms of its capacity to take advantage of the new emphasis that welfare reform put on paid employment.

Under-Reporting of Means-Tested Transfer Benefits in the CPS

❖Laura Wheaton, The Urban Institute, 2100 M Street NW, Washington, DC 20037, *lwheaton@ui.urban.org*

Key Words: CPS, underreporting, microsimulation, welfare, SIPP, ASEC

Underreporting of means-tested benefits in the Annual Social and Economic Supplement to the Current Population Survey (CPS) has increased since the early 1990s. This presentation will show trends in underreporting of SSI, AFDC/TANF, food stamps, and Medicaid/SCHIP from 1993 to 2005 and discuss the extent of imputation for non-response on the CPS. Comparisons will be made to the Survey of Income and Program Participation (SIPP) for selected years. Results from the 2004 current law simulation of the TRIM3 microsimulation model will be presented to demonstrate the use of microsimulation in correcting for underreporting of means-tested benefits in the CPS.

Dial M for Missing Murder Data

❖James A. Fox, Northeastern University, College of Criminal Justice, Boston, MA 02115, *j.fox@neu.edu*; Marc L. Swatt, Northeastern University

Key Words: Imputation, Muliple Imputation, Homicide

The FBI's Supplementary Homicide Reports represent the most valuable source of information on patterns and trends in murder and non-negligent manslaughter. Despite widespread use by researchers and policy-makers, these data are not without their limitations—most notably, missing or incomplete incident reports. In this analysis, we invoke methods for addressing missing data in the 1976-2005 SHR cumulative file, related to both non-reports (unit missingness) and incomplete reports (item missingness). For missing characteristics on victims, offenders or incidents, we apply a multiple imputation approach based on Joseph Schafer's (1997) algorithm for incomplete multivariate categorical data. Then, to handle unit missingness, we develop a weighting scheme linking FBI estimates of state homicide counts with NCHS mortality data on decedent characteristics in coroners' reports for homicides.

Measurement and Evaluation of Ancestry: Differences Between Specific and Generalized Responses

❖Frank Hobbs, U.S. Census Bureau, Population Division, Washington, DC 20233-8800, *frank.b.hobbs.iii@census.gov*; Angela Brittingham, U.S. Census Bureau

Key Words: ancestry, ethnicity, identity, Census 2000, American Community Survey

General responses to identity-related questions in U.S. censuses and surveys frequently are interpreted as less than desirable, and data collection and processing procedures may be modified to elicit more detailed categorizations. This paper uses Census 2000 and American Community Survey (ACS) data to evaluate differences between the populations reporting specific versus generalized responses to the questions on ancestry and Hispanic origin. Particular focus is on those who provide American as their ancestry and on those who provide Hispanic or Latino as their response to the question on Hispanic origin. People who do not provide an ancestry response also are examined. Do their characteristics substantively differ from those providing specific responses? Geographic patterns of general responses and differences by age, race, nativity status, and language spoken at home are investigated.

Measuring Religious Change Among Michigan Southern Baptists

❖William Livingston, Baker College, Center for Graduate Studie, 1116 West Bristol Road, Flint, MI 48507-5518, *bill.livingston@baker.edu*

Key Words: religious change, analysis of ordinal data, changing religious attitudes, statistical analysis of religious change

Sociologists have always been interested in analyzing and measuring social change. The statistical analysis of changing religious attitudes brings special challenges because of the changing definition of variables from generation to generation. This presentation analyzes and measures changing attitudes in a Michigan religious group, drawing data from three studies completed by this author over the most recent twenty year period.

256 ASA College Stat Bowl I

The ASA, IMS, ENAR, WNAR, SSC

Tuesday, July 31, 10:30 am–12:20 pm

Game 1

❖Ellen Breazel, University of Georgia; ❖Ann Lazar, University of Colorado at Denver and Health Sciences Center; ❖Chris Rigdon, Arizona State University; ❖Tanya Sikdar, University of Georgia

Game 2

❖Ming-Hung (Jason) Kao, University of Georgia; ❖Amanda-Sue King, Wright State University; ❖Peng Liu, Case Western Reserve University; ❖Balaji Raman, University of Connecticut

Game 3

❖Andrejus Parfionovas, Utah State University; ❖Sourish Das, University of Connecticut; ❖Yiying (Richard) Fan, Case Western Reserve University; ❖Qin Wang, University of Georgia

Game 4

❖Jien Chen, University of Georgia; ❖Xiaosong Li, Case Western Reserve University; ❖Sourish Saha, University of Florida; ❖Erin Tewksbury, Wright State University

257 Recent Advancements in Nonparametric Methods for Regression Model Selection and Model Checking ✪

Section on Nonparametric Statistics, Biometrics Section

Tuesday, July 31, 10:30 am–12:20 pm

Recent Developments in Noise-Addition Methods for Model Selection

❖Dennis Boos, North Carolina State University, Dept. of Statistics, Box 8203, Raleigh, NC 27695-8203, *boos@stat.ncsu.edu*; Leonard A. Stefanski, North Carolina State University

Key Words: False Selection Rate, Variable Selection, Regression

Luo, Stefanski, and Boos (2006, Technometrics) introduced Noise Addition Model Selection (NAMS), and Wu, Boos, and Stefanski (2007, JASA) introduced Variable Addition Model Selection (VAMS), both in the context of linear regression. This talk will update advances in these approaches along with extensions to a variety of situations.

Frequentist-Bayes Lack-of-Fit Tests Based on a Laplace Approximation

❖Jeffrey Hart, Texas A&M University, Department of Statistics, College Station, TX 77843, *hart@stat.tamu.edu*

Key Words: Bayes tests, Lack-of-fit tests, Nonparametric tests, Orthogonal series

The no-effect hypothesis in regression is tested in frequentist fashion using as test statistic a Laplace approximation to the posterior probability of this hypothesis. Dependence of the Laplace approximation on prior probabilities allows the investigator to tailor the test to particular alternatives. Use of diffuse priors produces new omnibus lack-of-fit statistics. These omnibus statistics are weighted sums of exponentiated squared (and normalized) Fourier coefficients, where the weights are model prior probabilities. Exponentiation of the Fourier components leads to tests that are much more powerful against higher frequency alternatives than classical tests, such as the cusum, based on weighted sums of squared Fourier coefficients. Our results appear to bring omnibus statistics that involve no choice of smoothing parameters back to the forefront of lack-of-fit testing.

A Goodness-of-Fit Test for Parametric and Semiparametric Models in Multiresponse Regression

❖Ingrid Van Keilegom, Université catholique de Louvain, Institute of Statistics, Voie du Roman Pays 20, Louvain-la-Neuve, International 1348 Belgium, *vankeilegom@stat.ucl.ac.be*; Song Xi Chen, Iowa State University

Key Words: bootstrap, empirical likelihood, goodness-of-fit, kernel estimation, monotone regression, partially linear regression

We propose an empirical likelihood test that is able to test the goodness-of-fit of a class of parametric and semiparametric multiresponse regression models. The class includes as special cases fully parametric models, semiparametric models, like the multi-index and the partially linear models, and models with shape constraints. It allows both the response variable and the covariate to be multivariate, which means that multiple regression curves can be tested simultaneously. The test also allows the presence of infinite dimensional nuisance functions in the model to be tested. It is shown that the empirical likelihood test statistic is asymptotically normally distributed under certain mild conditions and permits a wild bootstrap calibration. The test enjoys good power properties against departures from a hypothesized model within the class.

Robust Tests in Regression Models with Omnibus Alternatives and Bounded Influence

❖Lan Wang, The University of Minnesota, 385 Ford Hall, School of Statistics, 224 Church Street SE, Minneapolis, MN 55455, *lan@stat.umn.edu*

Key Words: robust tests, conditional moment test, influence function, lack-of-fit tests, smooth alternative, regression models

A robust approach for testing the parametric form of a regression function versus an omnibus alternative is introduced. This generalizes existing robust methods for testing subhypotheses in a regression model. The new test is motivated by developments in modern smoothing-based testing procedures and can be viewed as a robustification of a smoothing-based conditional moment test. It is asymptotically normal under both the null hypothesis and local alternatives. The robustified test retains the "omnibus" property of the corresponding smoothing test, i.e., it is consistent for any fixed smooth alternative in an infinite dimensional space. It is shown that the bias of the asymptotic level under shrinking local contamination is bounded only if the second-order Hampel's influence function is bounded.

258 Nonparametric Bayesian Inference for Challenging Biomedical Problems ● ✪

International Society of Bayesian Analysis, Section on Bayesian Statistical Science, ENAR, Biometrics Section, Section on Nonparametric Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Bayesian Semiparametric Factor Modeling in Large P Problems

❖ Mike West, Duke University, Box 90251, Durham, NC 27708-0251, *mw@stat.duke.edu*

Key Words: Bayesian factor models, Dirichlet process priors, Gene expression analysis, Biological pathway, Cancer genomics

Large-scale Bayesian latent factor models are now central tools in a range of our studies of complex, interconnected substructure in gene pathway biology. This talk will overview some of the models and methodology, and highlight the integration of non-parametric model components that very naturally reflect what can be highly non-Gaussian distributions in factor space. Examples drawn from cancer pathway studies will highlight the methodology.

The Matrix Stick-Breaking Process: Flexible Bayes Meta-Analysis

❖ David B. Dunson, National Institute of Environmental Health Sciences, NIH, P.O. Box 12233, MD A3-03, 111 T.W. Alexander Dr, Biostatistics Branch, Research Triangle Park, NC 27709, *dunson1@niehs.nih.gov*

Key Words: Meta analysis, Random effects, Hierarchical model, Bayesian, Nonparametric, Dirichlet process

In analyzing data from multiple related studies, it is often of interest to borrow information across studies and to cluster similar studies. Although parametric hierarchical models are commonly used, a concern is sensitivity to the form chosen for the random effects distribution. A Dirichlet process (DP) prior can allow the distribution to be unknown, while clustering studies. However, the DP does not allow local clustering of studies with respect to a subset of the coefficients without making independence assumptions. Motivated by this problem, we propose a matrix stick-breaking process (MSBP) as a prior for a matrix of random probability measures. Properties are considered, computational methods are developed, and the approach is illustrated using data from a multinational uterotrophic bioassay study.

Bayesian Nonproportional Hazards Survival Analysis

❖ Wesley Johnson, University of California, Irvine, Department of Statistics, Irvine, CA 92697, *wjohnson@uci.edu*

Key Words: Nonparametric, Gibbs Sampling, Dependent Dirichlet Process

We discuss some recent advances in nonparametric approaches to survival data. In particular we discuss a model that allows survival curves to cross, possibly due to a toxic treatment effect resulting in early deaths but which is ultimately effective as a treatment. An illustration based on a cancer clinical trial is given. As time permits, we also discuss some issues related to model selection for survival data with time dependent covariates.

259 The National Children's Study: The Challenges and Promises of a Follow-Up Study of a Randomly Selected Cohort of 100,000 Pre-and Post-Conception Women and Their Offspring through 21 Years of Life ●

ENAR, Biometrics Section, Section on Statistics and the Environment, Section on Statistics in Epidemiology, Section on Survey Research Methods, Section on Health Policy Statistics, Social Statistics Section, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 10:30 am–12:20 pm

The National Children's Study (NCS): Background, Rationale, and Scope

❖ Peter C. Scheidt, National Institute of Child Health and Human Development, NIH, 6100 Executive Blvd., Room 5C01, MSC 7510, Bethesda, MD 20892, *scheidtp@nih.gov*

Key Words: longitudinal study, cohort study

To address growing concerns about the vulnerability of children to environmental exposures and developmental and health conditions, Congress directed NICHD, with a consortium of Federal agencies (EPA, CDC, NIEHS), to conduct a longitudinal study of environmental influences on child health and development. 100,000 women are to be enrolled during pregnancy and their children studied through age 21. A national probability sample based on frequency of live births was chosen as the appropriate sample for the Study. 105 locations (counties) comprise the first stage of the multistage sample that will collect exposure, outcome and genetic data. 7 locations and Centers along with a Coordinating Center have been selected to begin the Study. When fully implemented, the NCS will be the largest in-depth longitudinal study of children's health, and a national resource for future investigations.

Methods for Sampling Households To Identify 100,000 Births

❖ Jill M. Montaquila, Westat, 1650 Research Blvd., Rockville, MD 20850, *montaqj1@westat.com*; J. Michael Brick, Westat; Lester Curtin, Centers for Disease Control and Prevention

Key Words: National Children's Study, longitudinal study

The National Children's Study is a national household probability sample designed to identify 100,000 children at birth and follow the sampled children for 21 years. Data from the study will support examining numerous hypotheses concerning genetic and environmental effects on the health and development of children. The goals of the study present substantial challenges. For example, the need for preconception, prenatal, and postnatal data require identifying women in the early stages of pregnancy, the collection of many types of data, and the retention of the children over time. In this paper, we give an overview of the sample design for this study,

and highlight the approaches used to address these challenges. We will also describe the rationale for the sampling choices made at each stage, the unique organizational structure of the NCS, and issues we expect to face during implementation.

Adaptive Tuning of Field Performance in the National Children's Study: Protection of the Inferential Base

❖Jonas H. Ellenberg, University of Pennsylvania, School of Medicine Center for Clinical Epidemiology and Biostatistics, 423 Guardian Drive, Blockley Hall #604, Philadelphia, PA 19104, *jellenbe@cceb.med.upenn.edu*

Key Words: National Children's Study, Longitudinal study, Missing data, field operations, Inference

The National Children's Study (NCS) is a national household probability sample designed to identify 100,000 pregnancies and follow the resulting offspring for 21 years. The inferential validity of the NCS will depend heavily on the field operations performance. Real time performance adjustments may include increased tracking of lost participants and increased personnel/training for biological specimen collection. In particular, providing prophylaxis rather than analytic treatment for missing data will be critical. Simultaneous longitudinal monitoring of performance metrics will provide a basis for realignment of resources to address performance. The approach, an interactive economic model constrained by a fixed budget, considers intervention costs, inferential validity and precision, and impact of NCS study results on the reduction of the US burden of disease.

Issues in the Measurement and Assessment of Environmental Exposures for a Large Prospective Cohort Study

❖James Quackenboss, U.S. Environmental Protection Agency, Div. of Epidemiology, Stats & Prevention, 6100 Executive Blvd., RM. 5C01, MSC 7510, Bethesda, MD 20892-7510, *quackenj@mail.nih.gov*; Warren Galke, National Institute of Child Health and Human Development; Susan Viet, Westat; David Marker, Westat; Steven Bodesky, LFR, Inc.

Key Words: environmental epidemiology, longitudinal study, exposure assessment, children's health

The National Children's Study (NCS) is a large-scale study that offers unique opportunities to evaluate relationships among multiple exposures and changes in susceptibility over life stages, and subsequent health and developmental outcomes. Exposure measures are selected based on Study hypotheses, and include chemical, physical, biological, and social environments. Given multiple hypotheses, these measures will be primary exposures in some instances and key covariates for others. Statistical and logistical considerations for exposure assessment designs and approaches will be discussed, including: integrating different types of exposure information, targeting assessments to key exposure pathways and locations, combining individual- and aggregate-level (e.g., community) information, and using nested case-control and other designs to minimize participant burden and costs.

260 Design of Experiment for Categorical Data ● ✪

Biometrics Section, ENAR, Biopharmaceutical Section, General Methodology, Section on Physical and Engineering Sciences, Section on Quality and Productivity, WNAR

Tuesday, July 31, 10:30 am–12:20 pm

A Response-Driven Design: The Randomized Reinforcement Urn

❖Nancy Flournoy, University of Missouri-Columbia, 146 Middlebush Bldg, Columbia, MO 65203, *flournoyn@missouri.edu*

Key Words: adaptive design, ethical treatment, clinical trials, asymptotic normality, optimal allocation, stochastic processes

Urn models are useful when randomization is important. Treatments assignments correspond to the color of a drawn ball. If balls are added or taken away depending on outcomes, the urn is said to be response-driven. Much work has focused on replacement procedures that increase the chances of allocating better treatments. The Randomized Reinforcement Urn (RRU) is optimal in the sense that it assigns patients to the best treatment with probability converging to one. No other urn model in the literature has this property. For binary random variables, this urn was studied by Durham, Flournoy and Li (1998). However, the allocation proportions degenerate to 0 or 1 (or to a random variable if treatments are equal) and theory that permits testing for treatment differences has only just been obtained (May and Flournoy, manuscript 2006). I will describe the general RRU and its application.

Optimal Designs for Dose-Finding Studies

❖Holger Dette, Ruhr-Universitaet Bochum, Fakultaet fuer Mathematik, Bochum, 44780 Germany, *holger.dette@rub.de*

Key Words: optimal design, doese finding experiments, robustness

Identifying the "right" dose is a very critical step in the clinical development process of any medicinal drug. Selecting too high a dose can result in unacceptable toxicity and associated safety problems, while choosing too low a dose leads to smaller chances of showing sufficient effectiveness in confirmatory trials. In this paper we consider the problem of obtaining dose finding designs, efficiently learning about the minimum effective dose (MED), defined as the smallest dose which shows a clinically relevant and a statistically significant effect. The sensitivity of these designs with respect to misspecification in the true parameter and in the underlying dose-response model is investigated. Robust optimal designs are constructed taking into account, simultaneously, a set of potential dose response profiles within classes of models commonly used.

Multicenter Clinical Trials with Binary Outcomes and Random Effects for Centers: Sample Size Determination

❖Kathryn Chaloner, The University of Iowa, Dept of Biotatistics, Dept of Statistics and Actuarial Science, 200 Hawkins Drive C22 GH, Iowa City, IA 52242, *kathryn-chaloner@uiowa.edu*

Key Words: Bayesian, Clinical trial, Sample Size

The question of choosing a sample size for a multi-center clinical trial where the analysis involves a binary response will be reviewed. A Bayesian hierarchical model is assumed. Centers are assumed to be exchangeable and so random effects for center are incorporated into the statistical analysis plan. Enrollment at each center is anticipated to be different, with some centers having much larger sample sizes than others. Of interest is not only to determine overall efficacy of an intervention, but also to require each center to demonstrate sufficient evidence of efficacy. Examples will be given, along with simulations to examine the properties of different designs. This work is joint with M. Katherine Cowles, Emine O. Bayman, Qian Cicci Shi and Lixun Zhang (University of Iowa).

Optimal Designs of Clinical Studies for Multivariate Probit Model

❖Valeri V. Fedorov, GlaxoSmithKline, 1250 S. Collegeville Rd, Collegeville, PA 19426, *valeri.v.fedorov@gsk.com*

Key Words: Multivariate probit model, Clinical trials, Dose-response, Design of experiments

While the narrative is built for the case of two binary endpoint the generalization for the higher dimension is straightforward. The primary goal is to establish (estimate) the dose-response relationship. The traditional optimal experimental design is concerned mainly with collective ethics: one tries to gain as much information as possible without explicit concern about patients in the trial, i.e., doing what is best for future patients. In contrast, the currently popular procedures for dose-finding studies are mainly concerned with individual ethics: doing what is best for current patients in the trial. The proposed approach attempts to compromise between these two extremes: we maximize information under constraints on a penalty function, which depends on efficacy and toxicity of each dose entering a design. Locally optimal, minimax, Bayesian designs are discussed.

261 Asymptotic Decision Theory and Its Applications ✪

IMS, Section on Nonparametric Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Adaptive Nonparametric Density Estimation via the Root-Unroot Transform and Wavelet Block Threshholding

❖Lawrence D. Brown, University of Pennsylvania, Wharton School, Department of Statistics, 3730 Walnut St., Huntsman Hall #400, Philadelphia, PA 19104, *lbrown@wharton.upenn.edu*

Key Words: density estimation, nonparametric regression, wavelets, Poisson regression, root-unroot transform, adaptive estimation

Nonparametric density estimation has traditionally been treated separately from nonparametric regression. Here, we propose an approach that first transforms a density estimation problem into a nonparametric regression problem. The algorithm for this involves suitably binning the observations and then transforming the binned data counts via a carefully chosen square-root transformation. A wavelet block-threshholding rule is then used for the regression problem, and produces an estimated nonparametric regression function. Finally an adjusted un-root transform is applied to yield the final nonparametric density estimator. The procedure is easy to implement. It enjoys a high degree of asymptotic adaptivity and is shown in numerical examples to perform well for standard density estimation settings.

Asymptotic Equivalence for Nonparametric Regression Experiments with Random Design

❖Andrew Carter, University of California, Santa Barbara, Dept of Statistics, Santa Barbara, CA 93106-3110, *carter@pstat.ucsb.edu*

Key Words: asymptotic equivalence, nonparametric regression

We explore an extension of Brown and Low's (1996) result establishing the asymptotic equivalence of nonparametric regression experiments and the continuous Gaussian-process-with-drift experiment. Specifically, we consider the effect of an unknown design densities in the regression experiment. The inclusion of the additional parameter also induces changes in the limiting Gaussian experiments - they are now a hierarchical model with the information about the design as one component. We will demonstrate the connection between the nonparametric regression experiments and the hierarchical Gaussian models which suggests a connection to Gaussian deconvolution problems.

Estimating Convex Functions

❖Mark Low, University of Pennsylvania, The Wharton School, 400 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104-6340, *lowm@wharton.upenn.edu*

Key Words: Nonparametric Function estimation, Adaptaion, Convex

An adaptive estimator for estimating convex functions in nonparametric regression is given which adapts to each unknown convex function.

Sufficiency and the Preservation of Fisher Information

❖David Pollard, Yale University, Box 208290 Yale Station, New Haven, CT 06520, *david.pollard@yale.edu*

Key Words: Fisher information, DQM

The Fisher information (FI) for a set of data is never smaller than the FI for some function of the data, T(data), with equality at least when T is a sufficient statistic. Recently, Kagan and Shepp (2005 Amer. Statistician) gave an example of an insufficient statistic for which FI(data) = FI(T). This talk will explore the relationship between loss of FI and sufficiency, when expressed using Hellinger differentiability and Le Cam's theory for equivalence of experiments.

262 Seasonal Time Series ●✪

Business and Economics Statistics Section

Tuesday, July 31, 10:30 am–12:20 pm

Recasting X11 Seasonal Adjustment Filters into Smoothers

❖Tucker S. McElroy, U.S. Census Bureau, 4700 Silver Hill Road, Washington, DC 20233-9100, *tucker.s.mcelroy@census.gov*; Agustin Maravall, Bank of Spain

Key Words: ARIMA model, nonstationary time series, seasonal adjustment

Two unsolved problems with the X-11 seasonal adjustment methodology are the construction of standard errors and the handling of the boundaries. We adapt the ``implied model approach" of Kaiser and Maravall (2005) to achieve both objectives, while using two different assumptions on the given generic filter. Namely, we consider either that it is an output-matching filter or that it is a Wiener-Kolmogorov (WK) filter. For both approaches we lay out the explicit assumptions that are necessary, and the resulting algorithms for generating smoothers. The methods are then applied to the X11 seasonal adjustment, trend, and irregular filters. We then produce time-dependent signal extraction Mean Square Errors for the X11 estimates.

New Models for Seasonal Time Series and Seasonal Adjustment

❖David Findley, U.S. Census Bureau, DIR, Room 5K152B, Washington, DC 20233, *david.f.findley@census.gov*; Brian Monsell, U.S. Census Bureau; John Aston, Academia Sinica; Jyh-Ying Peng, Academia Sinica

Key Words: RegARIMA models, Trading day effects, X-12-ARIMA, X-13A-S

Models with regression mean functions and seasonal ARIMA disturbances are heavily used by the seasonal adjustment community. We will describe, and show empirical results for, a new regression model for trading day effects in inventory (stock) series and a new class of seasonal ARIMA models. The new regression model is a constrained form of the stock trading day model of Bell (1984) obtained by imposing constraints on the flow series trading day coefficients, a procedure for which we have developed general methods. The new ARIMA models decompose the seasonal moving average factor of Box-Jenkins seasonal models into a product of two or three factors, each with its own coefficient, which is associated with a subset of the frequencies, 0, 1, ., 6 cycles per year.

Quantifying and Measuring Revisions in Time-Series Estimates

❖Duncan Elliot, Office for National Statistics, Government Buildings, Cardiff Road, Newport, Gwent, International NP10 8XG United Kingdom, *duncan.elliott@ons.gov.uk*; Craig H. McLaren, Office for National Statistics; Xichaun (Mark) Zhang, Australian Bureau of Statistics

Key Words: revision, seasonal adjustment, trend, decomposition

Time-series estimates, in particular seasonally adjusted and trend estimates, are widely used for analysis. Confidence in these estimates is often related to the size of revisions to past data points. In official statistics, revisions occur for a number of different reasons from changes in methodology to established revision policies for the publishing of data. Information on revisions can vary from simple measurements such as mean absolute percentage error to breakdowns of the source of revisions in the original estimates. We explore methods for providing detailed information on revisions, particularly related to the seasonally adjusted and trend estimates. These approaches are illustrated with real data.

Empirical Evaluation of X-11 and Model-Based Seasonal Adjustment Methods

❖Stuart Scott, Bureau of Labor Statistics, 5611 Lee Hwy, Arlington, VA 22207, *scott.stuart@bls.gov*; Richard Tiller, Bureau of Labor Statistics; Daniel Chow, Bureau of Labor Statistics

Key Words: time series decomposition, ARIMA model, X-12-SEATS

X-11 and ARIMA model-based seasonal adjustments are compared for 82 series from three Bureau of Labor Statistics programs, 1) establishment employment, hours, and earnings, 2) consumer prices, and 3) producer prices. When both methods are applied automatically, the model-based approach exhibits more flexibility; its seasonal component can be virtually deterministic or rapidly varying. The model-based method tends to estimate a more stable seasonal component for series which are noisy or mildly seasonal. For some series, comparisons are made for "analyst adjustments," in which analysts attempt to improve on the automatic adjustments. Often "automatic" results from one method lead to improvements in analyst adjustments with the other method. With the current state of the art, it seems easier to improve X-11 adjustments.

263 Statistical Analysis and Forecast of Space-Time Wind Power Data ● ✪

IMS, Section on Statistics and the Environment

Tuesday, July 31, 10:30 am–12:20 pm

Calibrated Probabilistic Forecasting at the Stateline Wind Energy Center: The Regime-Switching Space-Time (RST) Method

❖Kristin A. Larson, 3TIER, 2001 6th Ave suite 2100, Seattle, WA 98121, *klarson@3tiergroup.com*; Tilmann Gneiting, University of Washington

Key Words: weather Prediction, Continuous ranked probability score, predictive distribution, spatio-temporal, truncated normal

With the global proliferation of wind power, accurate short-term forecasts of wind resources at wind energy sites are becoming paramount. Regime-switching space-time (RST) models merge meteorological and statistical expertise to obtain accurate and calibrated, fully probabilistic forecasts of wind speed and wind power. The model formulation is parsimonious, yet takes account of all the salient features of wind speed: alternating atmospheric regimes, temporal and spatial correlation, diurnal and seasonal nonstationarity, conditional heteroscedasticity, and non-Gaussianity. The RST method identifies forecast regimes at the wind energy site and fits a conditional predictive model for each regime. Geographically dispersed meteorological observations in the vicinity of the wind farm are used as off-site predictors.

Nonparametric Assessment of Properties of Space-Time Covariance Functions Associated with Wind Speed Data

❖Bo Li, National Center for Atmospheric Research, 1300 30th Street Apt B3 24, Boulder, CO 80303, *boli@ucar.edu*; Marc Genton, University of Geneva; Michael Sherman, Texas A&M University

Key Words: Asymptotic normality, Covariance, Full symmerty, Separability

We propose a unified framework for testing a variety of assumptions commonly made for covariance functions of stationary spatio-temporal random fields. The methodology is based on the asymptotic normality of space-time covariance estimators. We focus on tests for full symmetry and separability, but our framework naturally covers testing for isotropy and Taylor's hypothesis. Our test successfully detects the asymmetric and nonseparable features in two sets of wind speed data. We perform simulation experiments to evaluate our test and conclude that our method is reliable and powerful for assessing common assumptions on space-time covariance functions.

Models for Short-Term Wind Speed Prediction

❖Amanda S. Hering, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, *mhering@stat.tamu.edu*; Marc Genton, University of Geneva

Key Words: Spatial time series, Prediction, Circular variables, Space-time simulation

The regime-switching space-time model proposed by Gneiting et al. (2006) used to predict average hourly wind speed at a wind energy farm works efficiently but is specific to its geographic location. The model is regime-dependent, changing based on whether the wind is blowing to the east or west. We eliminate the regimes and include wind direction as a covariate in the model. The two hour forecasts improve upon Gneiting et al.'s (2006) best model. Comparing the two models leads us to simulate both wind speed and wind direction over space and time. Wind direction must be treated as a circular variable, whose values are close when observations are near the endpoints of its range, and this introduces additional complexity to the simulation. Of particular interest are characteristics of the data that would make Gneiting et al.'s (2006) model superior to our and vice versa.

264 Confronting the Realities of Massive Imputation in Complex Surveys ● ✪

Section on Survey Research Methods, Section on Health Policy Statistics, Section on Government Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Preservation of Skip Patterns and Covariance Structure through Semiparametric Whole-Questionnaire Imputation

❖David R. Judkins, Westat, 1650 Research Blvd., Rockville, MD 20850, *DavidJudkins@Westat.com*; Thomas Krenzke, Westat; Zizhong Fan, MedImmune, Inc.; Andrea R. Piesse, Westat

Key Words: Predictive mean matching, Unordered polytomous models

Hot-deck imputation schemes are attractive because of how well they preserve complex features of marginal distributions, such as heaping of income reports at round figures, in addition to marginal means and variances. Historically, they have been less successful at preserving multivariate structure. The authors have previously reported on recent methodology for preserving important features of the multivariate structure of an entire questionnaire, including skip rules and the strongest bivariate associations among ordered variables. This methodology is based on predictive mean matching with recursive unsupervised modeling of ordered and binary variables to be imputed. They have recently extended the methodology to include preservation of associations among unordered variables. They will report on this methodology, including a simulation study.

Imputation in a Multimode Multi-instrument Study of Cancer Care

❖Yulei He, Harvard Medical School, 180 Longwood Ave., Boston, MA 02115, *he@hcp.med.harvard.edu*; Alan M. Zaslavsky, Harvard University

Key Words: cancer, missing data, nonresponse weighting, sequential regression multiple impuatation, survey

The Cancer Care Outcomes Research and Surveillance (CanCORS) Consortium is a multisite, multimode and multiwave study examining the care delivered to population-based cohorts of newly diagnosed patients with lung and colorectal cancer and assessing predictors and outcomes of that care. Missing data are a serious concern for the CanCORS. We use weighting and multiple imputation to handle nonresponse in the baseline survey. The baseline survey uses several different instruments applicable to various groups of patients; consequently, the target population must be clearly defined for each analysis so we can calculate appropriate nonresponse weights. For imputation, we applied the sequential conditional regression imputation approach, which specifies a collection of models regressing incomplete outcomes on other covariates. We assess the performance of this approach in this complex dataset.

Multiple Imputation Analysis of Association Between Cardiovascular Disease Risk Factors and Environmental Exposures

❖Trivellore E. Raghunathan, University of Michigan, M4071 SPH II, 1420 Washington Heights, Ann Arbor, MI 48109-2029, *teraghu@umich.edu*; Wei Chen, Karmanos Cancer Institute

Key Words: Multiple Imputation, Time Series Analysis, Bayesian Methods, Thin Plate Splines

There is growing public health concern about the effect of exposure to airborne particles on cardiovascular disease. The key issue is, however, how to estimate exposure levels at residential addresses which are usually not available but several measures available at various monitoring sites. This paper considers the estimation of PM10 and PM2.5 at residential addresses using the data collected by the Environmental Protection Agency (EPA) from thousands of monitoring stations nationwide and using them to relate to cardiovascular disease and risk factors. Spatial effects are modeled using a nonparametric approach based on thin-plate splines. Time effects are modeled with trend, cyclical, autoregressive effects. Models include spatial covariates population density and environmental factors: temperature, visibility and TSP. Bayesian multiple imputation framework is used for the analysis.

265 Causal Inference: Past, Present, and Future ● ✪

Council of Chapters, Biometrics Section, General Methodology, WNAR

Tuesday, July 31, 10:30 am–12:20 pm

Statistics and Causal Inference: Past and Present

❖Paul Holland, Educational Testing Services, Rosedale Road, Mailstop 12-T, Princeton, NJ 08541, *PHolland@ETS.ORG*

Key Words: Neyman-Rubin Model, causal inference, regression coeficients, observational studies

The Neyman-Rubin (NR) model has become the basis for much of the statistical thinking about measuring causal effects and the causal interpretation of regression coefficients and related model parameters. This paper will review the NR model and its antecedents and emphasize how it is expanded to accommodate causal inference problems of increasing complexity.

The Mathematics of Causal Inference in Statistics

❖Judea Pearl, University of California, Los Angeles, Cognitive Systems Lab, 4532 Boelter Hall, Los Angeles, CA 90095, *judea@cs.ucla.edu*

Key Words: Neyman-Rubin Model, Nonparametric structural equations, Calculus of causation, Causal effect estimation, Instrumental variables, Integration of data

The Neyman-Rubin (NR) model, through which statisticians were first introduced to causal analysis, suffers from two drawbacks: (1) It lacks formal underpinning and (2) it uses an opaque, unnatural notation for expressing causal assumptions. I will describe the mathematical basis of the NR model using nonparametric structural equations, and will provide a transparent mathematical notation for expressing and discerning assumptions. The result is a complete and coherent calculus of causation that has resolved several classical problems of interest, including questions of confounding, covariate selection, causal effect identification, legal responsibility, effect decomposition, instrumental variables and the integration of data from diverse studies. Reference: J. Pearl, Causality (Cambridge University Press, 2000) *http://bayes.cs.ucla.edu/jp_home.html.*

Causal Inference as a Missing Data Problem: Past, Present, and Future

❖Donald B. Rubin, Harvard University, Statistics Department, Science Center, One Oxford St, Cambridge, MA 02138-2901, *rubin@stat.harvard.edu*

Key Words: Neyman-Rubin Model, Missing data, Potential outcomes

The problem of drawing inferences about causal effects has been conceptualized intuitively for many years as one of missing data, but has been formally treated this way only in the last century. The continuing development of this formalization has led to substantial insights into difficult problems, and it is anticipated that further insights will be forthcoming in the future. This presentation will give an idiosyncratic overview of this perspective.

266 Recent Developments of Statistics in Forest, Wildlife, Fisheries, and Ecological Research ● ✪

WNAR, ENAR, Biometrics Section, Section on Statistical Consulting, Section on Statistics and the Environment

Tuesday, July 31, 10:30 am–12:20 pm

Searching for Ecological-Environmental Relationships

Eric P. Smith, Virginia Polytechnic Institute and State University; ❖Samantha Bates Prins, James Madison University, Department of Mathematics and Statistics, 322 Roop Hall, MSC 1911, Harrisonburg, VA 22807, *prinssc@jmu.edu*

Key Words: environmetrics, clustering, stressors, voronoi, multivariate, redundancy analysis

Biological monitoring of aquatic sites is used for trend assessment, site assessment and for evaluation of the restoration of sites. Despite the vast amount of information that is collected the analysis of this data remains difficult, especially when attempting to assign causes to changes in biological measures. One approach that we have investigated is to search for strong relationships between biological data and potential stressors and covariates. We use a search algorithm that is based on Voronoi tessellations to find irregular shapes that contain a set of spatially connected sites. Regression and multivariate methods are used to describe the relationships between biological metrics and stressors or between counts of taxa and stressors. The methods are illustrated on data sets from West Virginia and on data from the mid-Atlantic Highlands.

Statistical and Nonstatistical Issues in Designing Large-Scale Monitoring Efforts for the Northern Spotted Owl

❖James A. Baldwin, USDA Forest Service, P.O. Box 245, Pacific Northwest Research Station, Berkeley, CA 94701, *jbaldwin@fs.fed.us*; Joseph Lint, USDI Bureau of Land Management; Martin G. Raphael, USDA Forest Service; Eric Forsman, USDA Forest Service; John A. Laurence, USDA Forest Service; Robert Anthony, U.S. Geological Survey

Key Words: spotted owl, monitoring, detectability

Designing large scale monitoring studies involves many more issues than the usual issues of statistical efficiency and power. We examine the efficiency and power of a proposed Northern Spotted Owl monitoring design along with some of the non-statistical issues such as budgets, access to remote sites, and (most important) what land managers can and cannot do with the results.

Bayesian Spatial Modeling of Vegetation Coverage Data in the Missouri Ozark Forest Ecosystem Project

❖Zhuoqiong (Chong) He, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65210, *hezh@missouri.edu*; Xiaoqian Sun, Clemson University

Key Words: Forest Ecosystem, Spatial Models, Markov Chain Monte carlo, Bayesian

This talk presents a new spatial model that takes account of the special data structure and treats a cluster of measurements as repeated measurements in one location. The model is applied to the analysis of the total vegetation coverage data in the Missouri Ozark Forest Ecosystem Project. An MCMC algorithm based on the shrinkage slice sampler is developed. The results show that the soil depth is an important factor while the aspect class is less important when modeling the total vegetation coverage. In addition, the strong spatial effect does exist in the data and four measurements in quadrats of a subplot are not strongly correlated but are not independent. Prediction of the total vegetation coverage at unmeasured locations is developed. Finally, possible generalizations are discussed.

Capture-Recapture Models: Uses of the Robust Design

❖Ken Pollock, North Carolina State University, Dept of Zoology, Raleigh, NC 27695-7617, *pollock@unity.ncsu.edu*

Key Words: capture-recapture, robust design, point counts, sampling, unequal catchability

The robust design which combines open and closed capture-recapture models in one analysis has now become a classic part of the capture-recapture literature. In this paper I will review uses of the robust design to allow for unequal catchability of animals, temporary emigration of animals, and separation of components of recruitment. I will then consider some recent work where we use the robust design to separate components of detection in avian point counts (probability of a bird singing and probability of detection given a bird sings). I will also consider possible errors in the detection history inherent in such problems and how to model them.

267 Computationally Intensive Methods

Section on Statistical Computing, ENAR, Section on Nonparametric Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Exploratory Statistical Software

❖Antony Unwin, University of Augsburg, MNF, Augsburg, 86135 Germany, *unwin@math.uni-augsburg.de*

Key Words: Exploratory, Software

Surprisingly, many statisticians have never used a spineplot and don't even know what it is. There are many other models and techniques that are not much used and not well known. This is partly because not everything can be taught and partly because not everything is needed by everybody. It is also partly because software implementing these features is not readily available or known. This talk discusses exploratory statistical software, what tools it should provide and how it should do it. Exploratory work needs fast and flexible support. The computer power needed for this is often underestimated though it influences how effective the software is. The tradeoff between additional functionality and superior performance is not usually considered. That may be why EDA is not much discussed and does not have high status, even though it is an essential component of any statistical study.

Reducing the Variability in Least Squares Cross-Validation Bandwidths

Jeffrey Hart, Texas A&M University; ❖Simon Sheather, Texas A&M University, Department of Statistics, College Station, TX 77843, *sheather@stat.tamu.edu*

Key Words: density estimation, bandwidth selection, cross validation

Least squares cross validation (LSCV) is a popular way of selecting bandwidths for density estimates. A major advantage of LSCV over other methods is that it is widely applicable and it performs better than its competitors when the true density has many modes. However, LSCV is highly variable in practice, and it can dramatically undersmooth, and as such lead to spurious bumpiness in the resulting density estimate. In this talk we show that is possible to reduce the variability in LSCV and to overcome the sporadic, but persistent, cases of dramatic under smoothing by LSCV.

K Models Clustering

❖James E. Gentle, George Mason University, 6051 Burkewood Way, Burke, VA 22015, *jgentle@gmu.edu*; Li Li, George Mason University

Key Words: clustering, statistical learning, model building

In many modeling applications, one model fits some observations well, while a different model fits other observations. The models that fit data in different regimes can be used to form clusters within the data. We view this situation similarly to that in which K means clustering is appropriate. While in the case of K means, the models in the clusters are simply means, we extend the clustering approach to more general models. Our objective is two-fold; we wish to identify clusters in the data, but we also wish to explore relationships of the variables to each other in each of the clusters. We investigate various approaches to this problem. All are computationally intensive. One method begins with groups based on K-means clustering, fits the best model in each from a catalog of models, and then iteratively moves outlying observations from one group to another and refits the models.

268 Influencing State-by-State Curriculum Standards in Statistics for K-12 Education ● ✪

ASA NCTM Joint Committee on Curriculum in Stats and Probability, Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 10:30 am–12:20 pm

Influencing State-by-State Curriculum Standards in Statistics for K–12 Education

❖Gail Burrill, Michigan State University, Division of Science and Mathematics Education, 116 N. Kedzie Lab, East Lansing, MI 48824-1316, *burrill@msu.edu*; ❖Christine Franklin, University of Georgia, Department of Statistics, 254 Statistics Building, Athens, GA 30602, *chris@stat.uga.edu*; ❖Katherine Halvorsen, Smith College, Dept of Mathematics and Statistics, Burton Hall 313, Northampton, MA 01063, *khalvors@email.smith.edu*; ❖Cathy Seeley, University of Texas at Austin, Austin Development Building, 2901 N. IH-35, Suite 2.200, Austin, TX 78722-2348, *cseeley@mail.utexas.edu*

Key Words: K-12 statistics education, State standards, ASA GAISE guidelines, College Board Standards, NCTM Standards, NCTM Curriculum Focal Points

Statisticians need to play a significant role in writing state standards for K through 12 mathematics and statistics because school curriculum decisions are determined by state standards. How can we have a say? Christine Franklin, leader of ASA's GAISE guidelines for K-12 education in statistics, will describe her experience in helping Georgia become a model state in this regard. Katherine Halvorsen, a member of the College Board Standards writing team, will describe the process of creating the Mathematics and Statistics College Board Standards. Two former NCTM presidents who are committed to statistics education, Cathy Seeley and Gail Burrill, will discuss the past, present, and future of statistics in the NCTM Standards and Curriculum Focal Points, ASA's part in their development, their influence at the state level and how statisticians can get involved in that process.

269 Overarching Issues in Risk Analysis ● ✪

Section on Risk Analysis, Section on Statisticians in Defense and National Security

Tuesday, July 31, 10:30 am–12:20 pm

Overarching Issues in Risk Analysis

❖Duane Steffey, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *dsteffey@exponent.com*; ❖Bruce Craig, Purdue University, 150 North University Street, West Lafayette, IN 47907-2067, *bacraig@stat.purdue.edu*; ❖Robert Obenchain, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285-0001, *ochain@lilly.com;* Walter W. Piegorsch, The University of Arizona, Tucson, AZ 85721, *piegorsch@math.arizona.edu*

Key Words: data quality, exposure assessment, cost prediction, model validation, risk perception, multiple scales

Given the increasing diversification of research on risk analysis, and especially the expanding multiplicity of contexts in which risk analysis is being used, the panel will attempt to articulate broad statistical issues in risk analysis that span multiple contexts. Possible examples of such issues are: Data quality, which threatens all risk analyses, and which is typically not incorporated into uncertainty characterizations; Principled tradeoffs among data, modeling and computation; Exposure assessment, especially in emerging contexts such as food safety; Cost prediction–because there is associated "physics," more attention is devoted to predicting system behavior than predicting costs. The panel will inform the planned SAMSI 2007-08 program on Risk Analysis, Extreme Events and Decision Theory.

270 Phase II/III Clinical Trials: Combine or Not To Combine ● ✪

Biopharmaceutical Section, WNAR, ENAR, Biometrics Section, Section on Health Policy Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Multiplicity in Adaptive Designs

❖Willi Maurer, Novartis Pharmaceuticals AG, Novartis Pharma AG, WSJ27 1028, Basel, 4002 Switzerland, *willi.maurer@novartis.com*

Key Words: multiplicity, multiple comparisons, clinical trial, adaptive design, group sequential design, type I error control

Potential sources of multiplicity in adaptive seamless designs and their implications on the control of type I errors and decision making will be discussed. Repeated testing of the same hypotheses in a group sequential setting with early rejection or acceptance of hypotheses at interim analyses, testing of multiple initial hypotheses (e.g., multiple comparisons with a control), and change of design features (e.g., sample size) may need to be taken into account simultaneously. For each of these cases methods exist to account for multiplicity. Dealing with a combination of them within an adaptive trial often can be done in different ways. Options will be discussed together with their performance characteristics and possible operational restrictions.

How Much Evidence Is Enough?

❖Kuang-Kuo Lan, Johnson & Johnson PRD, 920 Route 202, Room 2671, Raritan, NJ 08869, *glan@prdus.jnj.com*

Key Words: adaptive design, inverse normal, inverse chi-square

Control of the alpha level is a key issue in drug development. Even though the traditional alpha level is set to be 0.025 (one-sided), it is extremely difficult to control the realized alpha at the desired level, and the power of an experiment will be affected accordingly. Combining evidence in Phase IIB/III studies is just one way to reduce the over-conservatism for alpha-control and improve power in drug development.

Adaptation of Adaptive Design in Japan

❖Masahiro Takeuchi, Kitasato University, 5-9-1, Shirokane, Minato-ku, Tokyo, 108-8641 Japan, *takeuchim@pharm.kitasato-u.ac.jp*; Hajime Uno, Kitasato University; Fumiaki Takahashi, Kitasato University

Key Words: global trial, bridging study

The US Food and Drug Administration published "Challenge and Opportunity on the Critical Path to New Medical Products" in 2004 in order to stimulate an innovative approach to the research and development of new medical products. It is well recognized that the excellent progress of basic science has discovered potential products, but that the slow progress of applied science has failed to deliver them to patients in a timely fashion. Therefore the concept of an adaptive design has been investigated extensively. The situation in Japan is very similar to the USA. Recently, the concept of the E5 guideline has been extended to a simultaneous global study through mutual extrapolation of clinical data among regions. We briefly overview the Japanese situation from perspectives of a clinical trial design and discuss the associated statistical issues.

Stagewise Planning for Clinical Trials from Phase II to Phase III

❖Sue-Jane Wang, Food and Drug Administration, 11903 New Hampshire Ave, CDER, Silver Spring, MD 20993, *suejane.wang@fda.hhs.gov*

Key Words: stagewise planning, adaptive design, flexible design, clinical trial, phase II, phase III

Traditionally, phase II trials are relatively small and can be expected to result in a large degree of uncertainty in the estimates based on which Phase III trials are planned. Phase II trials are also to explore dose regimens with high probability of showing effectiveness, appropriate primary efficacy endpoint(s) or patient populations. When the biology of the disease and pathophysiology of disease progression are well understood, the phase II and phase III studies may be performed in the same patient population with the same primary endpoint and the selected dose. In disease areas where the clinical outcome endpoint may not be observed in a short-term study (e.g., mortality in cancer) or the molecular pathways may not be well established, the treatment effect may be measured through an intermediate surrogate endpoint in phase II trials.

271 Using Computer Models for Ecological, Environmental, and Biological Applications: A SAMSI Program ● ✪

Section on Bayesian Statistical Science, ENAR
Tuesday, July 31, 10:30 am–12:20 pm

How Far Can a Physicist Travel Along the Bayesian Way: Probability, Uncertainty, and Adequacy in Climate-Like Modeling

❖Leonard Smith, London School of Economics, Houghton Street, London, WC2A 2AE United Kingdom, *lenny@maths.ox.ac.uk*

Key Words: model inadequacy, physical modelling, decision support, nonlinear

The roles of probability and uncertainty in physics based simulation modeling are considered within two hierarchies of physical models. It is argued that Bayesian notions are relevant only when state-of-the-art models are arguably empirically adequate. Given a model class no member of which accurately reflects the dynamics of the system under current conditions, how might one develop rational probabilistic beliefs regarding how the system would behave under altered conditions? What role does the probability calculus (including Bayes' Theorem) play in this situation? A slightly lower target than empirical adequacy, Probabilistic Similarity, is introduced. The current state of climate modeling, the rationality of many climate modelers, and the relevance of probabilistic analysis of current climate model output for decision support are each considered in this context.

Analyzing Input and Structural Uncertainty of a Hydrological Model with Stochastic Time-Dependent Parameters

❖Peter Reichert, Swiss Federal Institute of Aquatic Science and Technology, Eawag, POB 611, Duebendorf, 8600 Switzerland, *peter.reichert@eawag.ch*; Johanna Mieleitner, Eawag

Key Words: model structure deficit, time-dependent parameters, bias

A recently developed technique for analyzing the causes of bias in results of deterministic models is applied to a simple hydrological model. The application of this technique to a simple 7-parameter conceptual hydrological model demonstrates the power of the technique. The parameters (including additional parameters for input modification) have significantly different potential for bias reduction. Although there is a minor potential for improvement of the underlying deterministic model, most of the bias seems to be caused by random effects, particularly rainfall input uncertainty. Propagation of these errors through the model leads to strongly correlated model output errors because of the internal storage mechanisms in the model.

Analysis of Seasonal Epidemiological Parameters

❖Ariel Cintron-Arias, Statistical and Applied Mathematical Sciences Institute/North Carolina State University, 19 TW Alexander Dr, PO Box 14006, Research Triangle Park, NC 27709-4006, *ariel@samsi.info*; Peter Reichert, Swiss Federal Institute of Aquatic Science and Technology; Alun Lloyd, North Carolina State University; H. T. Banks, North Carolina State University

Key Words: seasons, communicable diseases, epidemiology, effective reproductive number, disease outbreaks

Seasonal fluctuations in the incidence of infectious diseases are a rather common phenomenon yet the underlying mechanisms have not been fully identified upon analysis of longitudinal observations. In the present study, seasonal epidemics are modeled by a differential equation model with periodic-time dependent transmission rates. Observed year to year variability in outbreak incidence can be explained by chaotic behavior of the solutions of such models or by additional variability in transmission rate parameters in time. To gain insight into possible causes for year to year variability, we investigate the temporal change of the transmission parameter required for a model without chaotic behavior that is required to reproduce measured data. This is compared to an analysis of a model with constant parameters showing chaotic behavior.

Hazard Mapping: Modeling Volcanic Avalanches

❖E. Bruce Pitman, University at Buffalo-SUNY, 318 Math Building, Buffalo, NY 14260, *pitman@buffalo.edu*

Key Words: volcanic avalanche, simulation, computer models, SAMSI

The risk of volcanic eruption is one that public safety officials regularly confront. Accompanying an eruption can be the hazardous flow of ash and rock that races along the mountain at speeds of tens of meters per second and runs out over distances of several kilometers. The physical processes activated in these flows are not well understood. However there are tools are under development to simulate volcanic avalanche flows. Investigation of the hazard posed by these flows must consider unknown or poorly constrained ingredients of the flow models, including variability of the flowing materials, the variable terrain over which a flow occurs, and the likelihood of an eruption event and its size. This talk describes the major issues confronting the scientific community in developing hazard maps for volcanic flows, and highlights recent progress towards solutions.

Calibration of an Air Quality Model

❖Serge Guillas, Georgia Institute of Technology, 686 Cherry Street, School of Math & Sciences, Atlanta, GA 30332-0160, *guillas@math.gatech.edu*; Lew Lefton, Georgia Institute of Technology; Yunsoo Choi, Georgia Institute of Technology; Yuhang Wang, Georgia Institute of Technology

Key Words: calibration, ozone, computer model, MCMC

We calibrate a regional chemical transport model (RAQAST) for surface ozone. The controllable input variables describe the meteorology. The calibration parameters are the diffusion (corresponding to boundary layer height), nitrogen oxides, anthropogenic VOC, biogenic isoprene, and effects of clouds on photolysis. A set of 100 runs of RAQAST enable us to find the best parameters. Time series and spatial approaches are discussed. We use a combination of Gaussian stochastic processes and MCMC techniques to find the best tuning parameters for forecasting purposes.

272 Genome-Wide Studies and Proteomics in Cancer ● ✪

Biometrics Section, ENAR, WNAR

Tuesday, July 31, 10:30 am–12:20 pm

Statistical Validation of Breast Cancer Subtypes

❖Amy Kapp, Stanford University, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305-4065, *akapp@stanford.edu*; Robert Tibshirani, Stanford University

Key Words: breast cancer, microarray, cluster analysis

For heterogeneous diseases, one of the most commonly used approaches to providing accurate and informative prognoses is to define subtypes. Subtypes traditionally have been defined using clinical parameters. The profusion of microarray data now available, however, makes it possible to define subtypes using molecular profiles. Based upon the idea of reproducibility, our method for validating the subtypes of a disease uses two independent microarray datasets and a new quantity called the in-group proportion (see the R package clusterRepro). The method is essentially a permutation test and several procedures for generating the null distribution of in-group proportions are discussed. When the method was applied to breast cancer tumor tissue microarray datasets using centroids previously defined, it provided evidence for the reproducibility of the basal, ERBB2+, and luminal B subtypes.

Analytical Pipeline for 2D Gels

❖Francoise Seillier-Moiseiwitsch, Georgetown University Medical Center, Suite 180 Building D, 4000 Reservoir Road NW, Washington, DC 20057, *seillier@georgetown.edu*

Key Words: proteomics, wavelets, 2D gels, alignment, images, electrophoresis

Two-dimensional polyacrylamide gel electrophoresis is a technique of choice to separate and display all proteins expressed in a tissue. In the resulting protein maps for groups of patients, we seek to identify differentially-expressed proteins. We describe a comprehensive analytical approach that deals with preprocessing, alignment and differential analysis. Preprocessing removes the bulk of the background noise. It involves smoothing, selecting regions containing spots and gradient thresholding. Images are aligned using cubic-spline transformations. The alignment is formulated as a quadratic programming problem that is optimized using an interior-point method. Wavelets are utilized to summarize the aligned images, and statistical tests are performed on the wavelet coefficients. These tests take into account the experimental design and have been developed with low sample sizes in mind.

Two-Stage Designs for Genome-Wide Association Studies in Cancer

❖Abhijit Dasgupta, Thomas Jefferson University, 1015 Chestnut St, Suite M100, Philadelphia, PA 19107, *adasgupt@mail.jci.tju.edu*

Key Words: Genome-wide association, SNP chip, Two-stage, linkage disequilibrium, breast cancer

The ongoing quest for putative genetic factors for cancer has been revolutionized by the arrival of high throughput SNP genotyping technologies and commercially available custom chips for SNP genotyping. These technologies have made large-scale genome-wide association studies possible, as well as opened up the possibility of fine-mapping regions of interest identified by the genome-wide studies. We will reconsider and extend the work of Satagopan and Elston (2003) and Wang et al (2006) to develop efficient, cost-effective two-stage designs for cancer studies that utilize genome-wide association to identify genomic regions of interest and subsequently genotype additional markers in these regions to identify regions of linkage disequilibrium with a disease SNP.

Effect of the Number of Top Genes on Survival Analysis in Microarray Data

❖Dung-Tsa Chen, Moffitt Cancer Center, University of South Florida, 12902 Magnolia Drive, Tampa, FL 33612, *dung-tsa.chen@moffitt.org*; Michael Schell, Moffitt Cancer Center; Steven Eschrich, Moffitt Cancer Center; Alan Cantor, Moffitt Cancer Center; Timothy Yeatman, Moffitt Cancer Center

Key Words: number of top genes, survival analysis, microarray

Gene expression profiling in cancer prediction often involves determination of a set of top genes for analysis, it is important to evaluate how modeling varies with the number of top genes incorporated. Our preliminary analysis indicates there is considerable variation of prediction outcome when the number of top genes changes. The variation implies that use of a different number K of top genes is likely to yield unreliable results. We propose a predictive risk probability approach to attempt to accommodate for the variation. This approach identifies a range from K to L top genes. From each number of top genes, the analysis identifies each patient as having high risk (score = 1) or low risk (score = 0). The categorizations are then averaged (majority vote), giving a high risk score between 0 and 1, thus providing a ranking for the patient's need for further treatment.

273 Adaptive Medical Device Clinical Trials ● ✪

Biopharmaceutical Section, ENAR

Tuesday, July 31, 10:30 am–12:20 pm

Comparison of Two Adaptive Randomization Methods in Multicenter Clinical Trials

❖ Lei Peng, Abbott Vascular, 3200 Lakeside Drive MS223, Santa Clara, CA 95054, *lpeng@guidant.com*

Key Words: Adaptive Randomization, Dynamic Randomization, Pocock-Simon

In small to moderate sized multi-center clinical trials, as the number of variables to be balanced increases, simple or stratified randomization could result in serious treatment imbalances. Adaptive randomization (AR) may be used. When both baseline variables and study centers need to be considered, the question arises as how should centers be handled in AR. Two AR methods are investigated: Pocock-Simon (PS) and Pocock-Simon-Zelen (PSZ). In PS, study center is treated similarly as baseline variables. In PSZ, baseline variables are initially used to make provisional treatment assignment using PS method. Treatment balances are then checked against study centers using Zelen's algorithm. Computer simulated data and chi-square goodness-of-fit test are used to compare the methods. Results show that PS method achieved better treatment balances with respect to baseline variables and centers.

Data Analysis in a Medical Device Trial When Planned Randomized Enrollment Is Too Slow

❖ Yonghong Gao, Food and Drug Administration, 1350 Piccard Dr, Rockville, MD 20850, *yonghong.gao@fda.hhs.gov*; Gerry Gray, Food and Drug Administration

Key Words: randomization, adaptation, medical device trial

In some medical device trials, the device under investigation is the first of a kind and "best medical therapy" is used as a control. If "best medical therapy" is not perceived by patients to be very effective, then they may be reluctant to enroll in a randomized clinical trial out of fear of randomization into the control arm. Hence the planned randomized enrollment is too slow to finish within a reasonable time frame. A possible mid-course adaptation is to add some non-randomized sites in the trial, allowing investigators to enroll and assign patients to treatment arm using their medical judgment. This will result to two different patient cohorts in the trial, a randomized cohort and a nonrandomized cohort. In this talk, different approaches to analyzing the data from the two cohorts are proposed and compared, and some simulation results are presented.

Statistical Analysis for Two-Stage Adaptive Design with Different Study Endpoints

❖ Shein-Chung Chow, Duke University, School of Medicine, 2400 Pratt Street, Room 0311 Terrace Level, Durham, NC 27705, *sheinchung.chow@duke.edu*; Siu-Keung Tse, Hong Kong City University

Key Words: Biomarker, Two-stage adaptive design, Sample size calculation, Different study endpoints

In the pharmaceutical industry, it is desirable to apply an adaptive seamless trial design to combine two separate clinical studies that are normally conducted for achieving separate objectives such as a phase II study for dose finding and a phase III confirmatory study for efficacy. In some cases, however, the study endpoints for the two separate studies may be different due to long treatment duration. In this case, test statistics for the final analysis based on the combined data are necessary developed. In this paper, a test statistic utilizing data collected from both phases is proposed assuming that there is a well established relationship between the two different study endpoints. Formula for sample size calculation based on the proposed test statistic is derived. Sample size allocation at the two phases is also discussed.

One Adaptive Design and Its Application to Medical Device PMA Study

❖ Xiaolong Luo, Johnson & Johnson PRD, 45 Technology Dr, Warren, NJ 07059, *xluo@crdus.jnj.com*; David Snead, Cordis, a Johnson & Johnson Company; Dennis Donohoe, Cordis, a Johnson & Johnson Company

Key Words: adaptive design, medical device, noninferority, spending function

Adaptive clinical trial design and analysis have been intensively studied and evaluated for pharmaceutical drug development during the past several years. It has the potential to shorten the study duration and minimize the patient resource. At the same time, it presents unique challenge to clinical operations and appropriateness of statistical analysis. There is additional challenge in its application in medical device area where double blind randomized control study may not be feasible in many situations. However, the need for optimizing trial design to better protect patients and to expedite the therapeutic options to medical community is equally important. In this talk, we will discuss some general concepts in adaptive design.

Flexible Designs in Clinical Trials: Case studies

❖ Ning Li, Food and Drug Administration, Rockville, MD 20850, *ning.li@fda.hhs.gov*

Key Words: Felxible designs, Clinical trials

The standard approach to design and analysis of confirmative clinical studies is to pre-specify all design features, including the primary endpoints and the test statistics. However, in flexible trial design, some flexibility to modify the pre-specified design features during the course of study is offered. Consequently, the trial sponsors may face variety of challenges to maintain the integrity and credibility of trials performed. In this talk, several real-scenarios cases of adaptive designed trials will be presented. The implicit and explicit impacts of the modifications on trial results/interpretations will be discussed.

274 Statistical Issues of a Thorough QT/QTc Study ● ✪

Biopharmaceutical Section

Tuesday, July 31, 10:30 am–12:20 pm

Effect of Baseline Measurement on the Change from Baseline in QTc Intervals

❖ Hong Tian, Johnson & Johnson PRD, 1000 Route 202, Raritan, NJ 08869, *htian@prdus.jnj.com*; Jaya Natarajan, Johnson & Johnson PRD

Key Words: QTc intervals, Baseline, Change from baseline, Parallel group designs

Phase I thorough QT (TQT) studies are routinely conducted by pharmaceutical companies for all new compounds to satisfy the requirements of ICH E14 guidance on the evaluation of QTc prolongation. The primary

endpoint is the change from baseline in QTc and the hypothesis of interest is the non-inferiority of drug to placebo. Sometimes, due to the properties of the compound, it becomes necessary to use parallel group designs for TQT studies. In such situations, the effect of the baseline on the change from baseline in QTc becomes an important issue since differing baseline between the drug and placebo groups may not allow for proper estimation of the drug's effect. In this work, we evaluate the effect of baseline on the change from baseline using the placebo data from several TQT studies. Resampling techniques are used to evaluate the impact of differing baselines across groups.

QT Interval Variability and Sample Size Consideration for Thorough QT Studies

❖Lu Zhang, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *lu.zhang@lilly.com*; Alex Dmitrienko, Eli Lilly and Company

Key Words: QT interval, Thorough QT study, Sample size, Variability, TQT

Eli Lilly and Company has recently conducted four thorough QT studies. A meta-analysis of these trials was performed to estimate the variability of QT interval and calculate variance components related to minute-to-minute, day-to-day variability, etc. The obtained results were used to develop a series of sample size calculation formulas that be utilized in powering future thorough QT studies. The proposed sample size calculation approach enables clinical trial researchers to account for various features of their thorough QT study, including study design (parallel and crossover designs), single-, double-delta methods, patient population (male and/or female subjects) and the number of ECG replicates.

Examination of Concentration Effect Modeling of QT Interval from a Statistical Perspective

❖Brian Smith, Amgen Inc., One Amgen Center Drive, MS 38 3 B, Thousand Oaks, CA 91320, *brismith@amgen.com*

Key Words: QT interval, Concentration effect modeling, Intersection-Union Method

Concentration effect modeling has been proposed as a more powerful alternative to the currently frequently used intersection-union approach. This talk will examine both methods with an aim of showing when and if concentration effect modeling can be safely used to make decisions regarding a molecule's effect on QT interval.

Statistical Issues of a Thorough QT/QTc Study

❖Joanne Zhang, Food and Drug Administration, 2928 Marlow Farm Terr, Silver Spring, MD 20904, *joanne.zhang@fda.hhs.gov*

Key Words: thorough QT/QT study, statistical testing, assay sensitivity

QT/QTc prolongation may cause ventricular arrhythmias including ventricular fibrillation and Torsade de Pointes, which can be fatal; therefore, the current ICH E14 Guidance requests all sponsors submitting new drug applications to conduct at least one thorough QT/QTc study, normally early in clinical development after some knowledge of pharmacokinetics of the drug. It is recommended that a double blinded, randomized, crossover or parallel QT/QTc study should be considered in healthy volunteers (some exceptions). The goal of a 'thorough QT/QTc study' (TQT) is to test if the mean QT/QTc difference of the drug and placebo is no greater than 10 ms (non-inferiority margin). If yes, we might conclude that the drug does not have the QT prolongation effect. In this presentation, I will address some statistical issues in terms of the design and analysis of a TQT study.

275 Diverse Applications of Modern Time Series ✪

IMS, Business and Economics Statistics Section

Tuesday, July 31, 10:30 am–12:20 pm

Local Whittle Estimation of Fractional Integration for Nonlinear Processes

❖Xiaofeng Shao, University of Illinois, Department of Statistics, 104-A Illini Hall, 725 S. Wright Street, Champaign, IL 61820, *ombao@uiuc.edu*; Wei Biao Wu, The University of Chicago

Key Words: Local Whittle estimation, Long memory, Nonlinear time series, Spectral analysis

We study asymptotic properties of the local Whittle estimator of the long memory parameter for a wide class of fractionally integrated nonlinear time series models. The formulation allows the widely used FARIMA models with GARCH innovations of various forms and our asymptotic results provide a theoretical justification of the findings in simulations that the local Whittle estimator is robust to conditional heteroskedasticity. Additionally, our conditions are easily verifiable and are satisfied for many nonlinear time series models.

Minimax Detection of Structural Change Using Large Deviations

❖Chuan Goh, University of Toronto, Department of Economics, 150 St George St, Toronto, ON M5S 3G7 Canada, *goh@economics.utoronto.ca*

Key Words: structural change, large deviations, limits of experiments, aligned ranks, rank test

This paper proposes an asymptotically optimal testing procedure for structural change in a dynamic time series regression. The test is derived by embedding the model in a sequence of statistical experiments in the classical sense of Blackwell (1951) and Le Cam (1964) and applying the theory of large deviations (LD) to that sequence. This approach allows one to characterize asymptotically minimax tests for shifts in mean against global alternatives with respect to a Chernoff-type risk function concerned with the LD probabilities of both type-I and type-II error. The approach taken here explicitly avoids possibly ambiguous characterizations of inferential efficiency in the case where tests for structural change in dynamic models are analyzed in the usual framework in which alternatives converge to the null of parameter stability in large samples.

State-Space Discrimination and Clustering of Atmospheric Time Series Data Based on Kullback Information Measures

❖Joseph Cavanaugh, The University of Iowa, Dept of Biostatistics C22-GH, 200 Hawkins Dr, Iowa City, IA 52242-1009, *joe-cavanaugh@uiowa.edu*; Thomas Bengtsson, Alcatel-Lucent

Key Words: atmospheric science, classification, geostatistics, state-space process

We consider the problem of identifying geographically homogeneous regions based on similarities in the temporal dynamics of weather patterns. Two disparity measures are proposed and applied to cluster time series of observed monthly temperatures from locations across the state of Colorado. The two measures are based on state-space models, where the monthly temperature anomaly dynamics and seasonal component variation are structurally represented in the latent process. Application of these

measures produces clusters consistent with known atmospheric flow structures. Importantly, our procedures allow for discrimination and clustering according to the individual, unobserved structural components of a state-space process.

Inference for Vector Processes with Long Memory via the EM Algorithm

❖ Nalini Ravishanker, University of Connecticut, 215 Glenbrook Road, Storrs, CT 06269, *nalini.ravishanker@uconn.edu*; Jeffrey Pai, University of Manitoba

Key Words: EM algorithm, Financial derivatives, Long memory, VARFIMA

A framework for estimation of parameters in vector autoregressive fractionally integrated moving-average (VARFIMA) models via the EM algorithm is presented. It is well known that such processes capture both short-term correlation structure and long-range dependence characteristics of the individual series, as well as interdependence and feedback relationships between the series. Series with Gaussian and heavy-tailed innovations are considered. An application to modeling multivariate daily temperatures at selected measurement sites and its relation to pricing financial derivatives is presented.

Spline-Backfitted Kernel Smoothing of Additive Models in Time Series

❖ Li Wang, Michigan State University, Department of Statistics and Probability, East Lansing, MI 48824-1027, *wangli4@stt.msu.edu*; Lijian Yang, Michigan State University

Key Words: Bandwidths, B-spline, knots, Nadaraya-Watson estimator, nonparametric regression

Application of non- and semi parametric regression techniques to high-dimensional time series data have been hampered due to the lack of effective tools to address the 'curse of dimensionality'. Under rather weak conditions, we propose a spline-backfitted kernel estimator of the component functions for the nonlinear additive time series data that is both computationally expedient so it is usable for analyzing very high dimensional time series, and theoretically reliable so inference can be made on the component functions with confidence. Simulation experiments have provided strong evidence that corroborates with the asymptotic theory. Finally, the estimation procedure has been illustrated by a US unemployment rate example.

276 High-Dimensional Inference and Spatial Analysis ✪

IMS, Section on Nonparametric Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Profile-Kernel Likelihood Inference with Diverging Number of Parameters

❖ Clifford Lam, Princeton University, Department of Operations Res and Fin Eng, Princeton, NJ 08544, *wlam@princeton.edu*; Jianqing Fan, Princeton University

Key Words: Generalized linear models, Varying coefficients, High dimensionality, Asymptotic normality, Profile likelihood, Generalized likelihood ratio tests

We investigate both theoretical and practical sides of profile likelihood estimation and inference for the generalized varying coefficient partially linear model with growing number of predictors. When the number of parameters grows with sample size, the existence and asymptotic normality of the profile likelihood estimator are established under some regularity conditions. Profile likelihood ratio inference for the growing number of parameters is proposed and Wilk's phenomenon is demonstrated. A new algorithm for computing profile-kernel estimator is proposed and investigated. Simulation studies show that the resulting estimates are as efficient as the fully iterative profile-kernel estimates. For moderate sample sizes, our proposed procedure saves much computational time over the fully iterative profile-kernel one and gives more stable estimates. A set of real data is also analyzed.

High-Dimension, Low Sample Size (HDLSS) Data Asymptotics

❖ Xuxin Liu, The University of North Carolina at Chapel Hill, Dept. of Statistics and Operations Research, 207 Conner Drive Apt 12, Chapel Hill, NC 27514, *liux@email.unc.edu*

Key Words: High Dimensional data, Microarray, Batch adjustment

We present the HDLSS asymptotics, in the sense that the dimension goes to infinity and the sample sizes are fixed. The geometric structure of the data cloud becomes very simple under HDLSS asymptotics. We compare two micro-array data combination methods, PAM and DWD. These two methods have very different HDLSS asymptotic properties. We give the conditions for the consistency and inconsistency of these two methods for the combination of two data sets with unbalanced sample sizes.

Model Selection Using Hellinger Distance: Methods and Applications

❖ Xiaofan Cao, Colorado State University, Campus Delivery 1877, Department of Statistics, Fort Collins, CO 80523-1877, *cao@stat.colostate.edu*; Haonan Wang, Colorado State University; Hariharan Iyer, Colorado State University

Key Words: Model Selection, Hellinger Distance, False Discovery Rate (FDR)

A model selection strategy based on an estimator of the expected Hellinger distance between an approximating model and the unknown true model is developed. Convergence and invariance properties of the proposed estimator are studied. The performance of the proposed model selection strategy is assessed in mixture distribution problems and in factorial ANOVA model selection problems using a statistical simulation study. The proposed model selection method and minimum Hellinger distance estimation are applied to the problem of estimation of the false discovery rate (FDR) in Microarray data analysis. The methods are illustrated using real data examples. The R codes for their implementation will be made available upon request.

Mixture of Thresholded, Transformed Gaussian Processes Modeling with Application to Rainfall Extremes

❖ Jie Zhou, The University of North Carolina at Chapel Hill, 210 Smith Building, CB3260, Chapel Hill, NC 27599, *jzhou@email.unc.edu*; Richard L. Smith, The University of North Carolina at Chapel Hill

Precipitation data are collected at point-source locations (rain gauges), or averaged over grid cells via a weather forecasting model or reanalysis. In previous work, a transformed Gaussian process model has been used to represent the spatial distribution of rainfall on any given day, to characterize the relationship between point-source and grid-averaged measurements. The current work extends this to a mixture of Gaussian processes, to allow for different weather patterns with different spatial characteristics of precipitation. A Monte Carlo procedure is proposed for estimating the model, that is then applied to US data both from rain gauges and from weather model reanalyses. The two data sources are compared by computing 50-year return values and provide an evaluation of the quality of reanalysis data for estimating extremes.

Sure Independence Screening for Ultra-High Dimensional Feature Space

❖Jinchi Lv, Princeton University, Department of Mathematics, 1106 Fine Hall, Princeton, NJ 08544, *jlv@princeton.edu*; Jianqing Fan, Princeton University

Key Words: Variable selection, dimensionality reduction, SIS, Dantzig selector, SCAD, adaptive Lasso

Variable selection is fundamental to high-dimensional statistical modeling. In a seminal paper, Candes and Tao (2007) propose a minimum l_1 estimator, the Dantzig selector, and show that it mimics the ideal risk within a logarithmic factor log p. Their procedure and result are challenged when dimension p is ultra high. In this paper we introduce the concept of sure screening and propose a fast method, Sure Independence Screening (SIS) to reduce high dimensionality to a relatively large scale, say below sample size. An appealing special case of SIS is the componentwise regression. In a fairly general asymptotic framework, SIS is shown to possess the sure screening property for even exponentially growing dimension. With ultra-high dimension reduced accurately to below sample size, variable selection can be accomplished by some refined lower-dimensional methods that have oracle properties.

277 When Rubber Meets the Road: Bayesian Clinical Trials in Action —MDA Experience

Section on Bayesian Statistical Science, ENAR

Tuesday, July 31, 10:30 am–12:30 pm

Bayesian Clinical Trials Designed and Conducted at the University of Texas M. D. Anderson Cancer Center (MDACC)

Swati Biswas, University of North Texas Health Science Center; ❖Diane D. Liu, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 447, Houston, TX 77030, *dianeliu@mdanderson.org*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center; Donald A. Berry, The University of Texas M.D. Anderson Cancer Center

Key Words: Bayesian Design, Clinical Trials, CRM, Adaptive Randomization

We reviewed 964 protocols registered in MDACC Protocol Document Online System between 2000–2005. Of these, 59% were MDACC only trials, 38% were multicenter trials and 3% were CCOP trials. The numbers of clinical trials by phases are: I: 106, I/II: 66, II: 348, II/III: 6, III: 78, IV: 7, and others: 353. Bayesian designs accounted for 20% of the trials reviewed. A majority of Bayesian designs (90%) were applied in the phase I and II settings. Most commonly used Bayesian designs conducted at MDACC include the continuous reassessment method (CRM), adaptive randomization, and designs to monitor efficacy and toxicity simultaneously. We will describe various types of innovative Bayesian trial designs that have been conducted at MDACC.

Five-Arm Dose Selection Trial with Adaptive Randomization

❖Mark F. Munsell, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 447, Houston, TX 77030, *mfmunsell@mdanderson.org*; Donald A. Berry, The University of Texas M.D. Anderson Cancer Center; J. Kyle Wathen, The University of Texas M.D. Anderson Cancer Center; Marcos J. de Lima, The University of Texas M.D. Anderson Cancer Center

Key Words: adaptive randomization, dose-selection

We describe a dose selection trial with a placebo arm and four active arms, each a different dose of the same agent. Doses too low or too high will not be effective. The study outcome is treatment success. Patients are at high or low risk of treatment failure. The study is planned with a maximum of 150 patients. Patients are randomized to the five arms, with patients assigned to placebo with probability 0.20. We find the probability that the treatment success rate in each active arm is greater than the treatment success rate in the placebo arm, while considering risk of failure. These probabilities aid in adapting the randomization probabilities for the four active arms, making decisions whether to close accrual to one or more arms and selecting the best active arm. We will discuss operating characteristics and practical considerations of the design.

An Application of Adaptive Randomization Using Hierarchical Bayes Model in a Prospective Biomarker-Based Clinical Trial

❖Suyu Liu, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd Unit 447, Houston, TX 77030, *syliu@mdanderson.org*; Edward S. Kim, The University of Texas M.D. Anderson Cancer Center; Xian Zhou, Genentech, Inc.; Ignacio Wistuba, The University of Texas M.D. Anderson Cancer Center; Roy Herbst, The University of Texas M.D. Anderson Cancer Center; Jeffrey Lewis, The University of Texas M.D. Anderson Cancer Center; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: adaptive randomization, clinical trial, hierarchical Bayes model, operating characteristics

To identify best-matched treatments for patients, we implement an outcome-based adaptive randomization in BATTLE trial (Biomarker-based Approaches of Targeted Therapy for Lung Cancer Elimination). It consists of 4 parallel phase II studies each with targeted therapies. Patients require a core biomarker biopsy prior to randomization. A hierarchical Bayes model is used to characterize efficacy rates among the 4 treatments for each biomarker profile. Based on the posterior clinical efficacy, patients are adaptively randomized according to their real-time biomarker status. The operating characteristics based on simulations indicate that the design can accurately identify the effective biomarker-treatment combinations, and allocates more patients to more efficacious treatments - a step toward “personalizing medicine”. Examples of data realization and practical considerations will be given.

Software and Practical Considerations for Designing and Conducting Innovative Clinical Trials at MDACC

❖John D. Cook, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd Unit 447, Houston, TX 77030, *jdcook@mdanderson.org*

Key Words: Bayesian clinical trials, adaptive randomization, dose-finding, safety monitoring

This presentation will survey the most commonly used software applications for designing and conducting Phase I and II clinical trials at M. D. Anderson Cancer Center. Practical considerations include explaining innovating methods to researchers, heuristics for selecting design parameters, and the logistics of patient accrual and attrition.

When Rubber Meets the Road: Bayesian Clinical Trials in Action—MDA Experience

❖ J. Jack Lee, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, *jjlee@mdanderson.org*

Key Words: Bayesian clinical trials, prior choice, software implementation, practical considerations

Although there have been no dearth of Bayesian clinical trial designs published in the literature, many of them are not put into practice. This session consists of four integrated talks showing Bayesian clinical trials in action at the M. D. Anderson Cancer Center. The types of trials, the design considerations, and their operating characteristics are presented. I will discuss key components for the successful implementation of these trials and major road blocks. I will also lead the discussion on the practical considerations and lessons learned in designing and conducting Bayesian trials. Audience participation is much encouraged.

278 Multi-State Life Table Methods in Health Policy Research ● ✪

Section on Health Policy Statistics

Tuesday, July 31, 10:30 am–12:20 pm

Evaluation of Using Life Table Methods Combined with Microsimulation To Produce Estimates of Distributions

James Lubitz, National Center for Health Statistics; ❖ Liming Cai, National Center for Health Statistics, 3311 Toledo Road, Rm 6330, Hyattsville, MD 20782, *lcai@cdc.gov*; Nathaniel Schenker, National Center for Health Statistics; Paula Diehr, University of Washington

Key Words: semi-Markov process, multi-state life table, healthy life expectancy, Aging

Life table methods have been widely used to estimate healthy life expectancy (HLE). But in almost all of the studies the estimates have been averages, not distributions. In this study, we apply life table methods and micro-simulation to the Cardiovascular Health Study (CHS) data to produce estimates of the distribution of the 65 and over population on years with functional limitations and years in fair or poor self-reported health states. We compare the distributions from the simulated population with the distribution of the actual sample on the same measures to provide insights into whether, and to what extent, simulated populations may be used to estimate population distributions. We estimate two life table models: the multistate life table model and the semi-Markov process model estimated using the stochastic EM algorithm (SMP-EM).

Age-Specific Prevalence and Years of Healthy Life in a System with Three Health States

❖ Paula Diehr, University of Washington, Box 357232, Seattle, WA 98195, *pdiehr@u.washington.edu*; David Yanez, University of Washington; Ann Derleth, University of Washington; Anne Newman, University of Pittsburgh

Key Words: multi-state lifetable, years of healthy life, self-rated health

Consider a three-state system with one absorbing state, such as Healthy, Sick, and Dead. Over time, the prevalence of the Healthy state will approach an “equilibrium” value that is independent of the initial conditions. We derived this equilibrium prevalence (Prev:Equil) as a function of the local transition probabilities. We then used Prev:Equil to estimate the expected number of years spent in the healthy state over time by adapting the Sullivan method for longitudinal data. This estimate is similar to one calculated by multistate life table methods and has the added advantage of having an associated standard error. In an example, we found that the standard error is accurate when a valid survival table is known from other sources, or when the available dataset is large enough to estimate survival accurately. It performs best when the number of waves of longitudinal data is not large.

Multistate Life Table Distributions for Highly Refined Subpopulations from Cross-Sectional Data: A Bayesian Alternative to Sullivan’s Method

❖ Scott Lynch, Princeton University, Department of Sociology, Wallace Hall 114, Princeton, NJ 08544, *slynch@princeton.edu*

Key Words: Gibbs sampling, multistate life tables, demography, Sullivan’s method

Sullivan’s method is the most commonly used method for estimating healthy life expectancy (HLE). Although the data requirements for the method are minimal, the method is limited in its ability to produce HLE estimates for subpopulations due to limited disaggregation of data in cross-sectional mortality files and small cell sizes in aggregated micro survey health data. We develop a method that adapts Sullivan’s approach to allow the inclusion of covariates in estimation of HLE and to generate interval estimates of HLE for any desired subpopulation that can be specified in the health data. The method involves (1) using Gibbs sampling to sample parameters from a hazard model; (2) using ecological inference to produce transition probability matrices from the Gibbs samples; and (3) applying standard multistate calculations to the transition probability matrices.

An Extension of the Cohort-Component Model of Population Projection: From Multistate Life Table to Multistate Model for Households and Living Arrangement Projections

❖ Kenneth Land, Duke University, Dept of Sociology, Durham, NC 27708-0088, *kland@soc.duke.edu*; Yi Zeng, Duke University

Key Words: Multstate Life Tables, Stochastic Forecasts, Demographic Models, Household Dynamics

We briefly summarize the multistate ProFamy modeling framework for household and population projections. We then focus on the rational, justification, current and potential applications of a unique, important demographic property of ProFamy model, namely, that it substantially extends the classic cohort-component model for projecting population age-gender distributions to projections that include not only age-gender distributions but also family households and living arrangements of all members of the population. Lastly, we describe current work on methodological issues pertaining to further extensions of the ProFamy model to include health/disability status of the elderly and to perform stochastic forecasting with confidence interval estimates.

Flowgraph Models for Multivariate Survival Data

❖ C. Lillian Yau, Tulane University, 1440 Canal St Suite 2001, New Orleans, LA 70112, *cyau@tulane.edu*

Key Words: flowgraph models, saddlepoint approximation, kidney transplant, semi-Markov models, multivariate survival data, multi-state models

Multistate statistical models describe multiple-event survival data that occur over time. They model stochastic processes that can progress through various states. The analysis focuses on modeling waiting times until events of interest occur. The theories are formalized within counting processes. While the theory is elegant, applications have been limited to Markov models. As an alternative, semi-Markov models have gained popularity recently. Among them, flowgraph models have demonstrated great flexibilities. Flowgraphs operate in moment generating function space, and allow virtually any waiting time distributions between transitions. Cumulative distribution function and probability density function are converted via saddlepoint approximations, The survivor and hazard are then computed. The method is illustrated using kidney transplant data.

279

Marketing Research Challenging Statistics: Problems Solved and Unsolved

Section on Statistics and Marketing

Tuesday, July 31, 10:30 am–12:20 pm

Methodology and Challenges in Search Marketing at Business.com

❖Nathan Janos, Business.Com, 2120 Colorado Blvd, Third Floor, Santa Monica, CA 90404, *njanos@business.com*; Tom Anderson, Business.Com; Paul Dagum, Business.Com

Key Words: search marketing, generalized linear mixed models, revenue optimization, profit optimization, portfolio, demand elasticity

Annual search marketing spending in the U.S. grew 28% to $7.9B in 2006 and is expected to double by 2010. This has lead to an increase in demand for search terms while posing new challenges in running a profitable campaign. Campaigns must systematically analyze large amounts of click data, adjust bid inferences across many search terms, maximize conversion and best monetize spend. At Business.com we have developed a demand-chain profit model for search marketing. We model demand elasticity using generalized linear mixed models with random cluster effects across groups of correlated search terms and repeated measures across search engine partners. Bid positions are adjusted daily using a portfolio profit and revenue optimization for more than 100,000 search terms. We discuss these techniques, information challenges and our results in managing a successful search marketing campaign.

Optimal Designs for Choice Experiments To measure the Willingness to Pay

❖Bart Vermeulen, Katholieke Universiteit Leuven, Naamsestraat 69, Leuven, 3000 Belgium, *bart.vermeulen@econ.kuleuven.be*; Peter Goos, Universiteit Antwesrpen; Martina Vandebroek, Katholieke Universiteit Leuven

Key Words: optimal design, willingness-to-pay, bayesian

The marginal rate of substitution (MRS) is the rate that measures the consumer's willingness to give up an attribute of a good in exchange for another attribute. An estimate of the MRS is usually obtained by taking the ratio of two parameter estimates associated with two explanatory variables of a discrete choice model. When price is one of the attributes, the MRS is called the willingness-to-pay (WTP). The use of choice experiments to calculate the WTP has received increasing attention in recent valuation literature. The design of the choice experiment determines the precision of the estimators of the parameters of the model, thus also of a WTP. In this work, we present a design optimality criterion to construct choice experiments which results in the most accurate estimators of a WTP. These designs are compared in terms of estimation accuracy to other commonly used design strategies.

Statistical Issues in Optimal Product Design

❖James Cochran, Louisiana Tech University, College of Business, PO Box 10318, Ruston, LA 71272, *jcochran@cab.latech.edu*

Key Words: Resampling, Marketing, Bias, Optimization, Political Science

Recent advances in algorithms and formulations enable us to solve product design problems to demonstrable optimality in a few seconds. Since sample data are used to estimate parameter values for these problems, the problems lie at the juncture of statistics and operations research; the maximands of these problems are estimates with associated sampling distributions and statistical characteristics. Our ability to rapidly solve these problems to optimality enables us to use resampling approaches that were heretofore computationally impractical to assess various statistical characteristics. We provide an overview of some product design problems that can now quickly be solved to optimality (including the single product design and political platform problems) and present approaches for assessing the variation and bias of sampling distributions associated with the maximands of such problems.

A Theory of Attribution

❖Barry Feldman, Prism Analytics/Russell Investment, 14637 Starr Rd SE, Olalla, WA 98359, *barry@prismanalytics.com*

Key Words: Attribtuion, cooperative games, market research, proportional value, relative importance, Shapley value

A general theory of attribution is developed. The probability that any ordering of factors in a general microeconomic model is properly ordered by relative importance is determined by elementary axioms. Two types of attribution result. These correspond to the Shapley (1953) value and the proportional value (Feldman (1999, 2002) and Ortmann (2000)) of a related cooperative game. Linear attribution based on the Shapley value is typically advocated as a decomposition method by marketing researchers (e.g., Conklin and Lipovetsky (2001) and Soofi, Retzer and Yasai-Ardekani (2000)). In proportional marginal attribution, the share of value received is equal to the probability that a factor is most important. The logic of linear and proportional marginal attribution is described. Examples highlight important differences between the linear and proportional marginal approaches.

Using Cooperative Game Theory as a Foundation for Marketing Research

❖Shon Magnan, GfK Custom Research North America, 8401 Golden Valley Rd, Minneapolis, MN 55427, *shon.magnan@gfk.com*; Stan Lipovetsky, GfK Custom Research North America; Ken Powaga, GfK Custom Research North America

Key Words: Game theory, marketing research, TURF, Customer Loyalty, Regressions

Much of marketing research is build upon a foundation of core statistical techniques. For example, in new product research, Total Unduplicated Reach and Frequency (or TURF) is often used in optimize product line decisions. In customer satisfaction, linear and additive models are often used to determine attribute importance in driving customer satisfaction. In data base modeling, standard regression model are used to build inferences about the connection between dependent and independent variables. Although these methods have long track records, each have fundamental flaws that impact their ability to produce actionable results. Faced with these problems, a direction of research has emerged that uses cooperative game theory and addresses many of these shortcomings. This presentation discusses these new approaches and how they are improving the quality of marketing research.

280 Optimal Designs: Ideas and Applications ● ✪

WNAR, ENAR

Tuesday, July 31, 10:30 am–12:20 pm

Optimal Allocation for Sequential Binary Response Experiments

❖ William Rosenberger, George Mason University, 4400 University Dr MS 4A7, Fairfax, VA 22030, *wrosenbe@gmu.edu*; Yevgen Tymofyeyev, Merck & Co., Inc.; Feifang Hu, University of Virginia

Key Words: adaptive designs, optimal allocation, sequential estimation, binary response

For sequential experiments with K treatments, we establish two formal optimization criteria to find optimal allocation strategies. Both criteria involve the sample sizes on each treatment and a concave noncentrality parameter from a multivariate test. We show these two criteria are equivalent. We apply this result to specific questions: how do we maximize power of a multivariate test of homogeneity with binary response? For fixed power, how do we minimize expected treatment failures? Because the solutions depend on unknown parameters, we describe a response-adaptive randomization procedure which ``targets” the optimal allocation and provides increases in power along the lines of 2–4 percent over complete randomization for equal allocation.

Optimal Allocation in Sequential Continuous Response Experiments

❖ Feifang Hu, University of Virginia, Department of Statistics, Charlottesville, VA 22904, *fh6e@virginia.edu*; Hongjian Zhu, University of Virginia

Key Words: clinical trials, doubly adaptive biased coin design, power, response-adaptive randomization

For sequential experiments with K treatments and continuous responses, we consider optimal allocations under following two formal optimization criteria: (i) maximize power of a multivariate test of homogeneity, and (ii) minimize the total expected response for a fixed power. The solutions depend on unknown parameters. We then use a response adaptive randomization procedure to target the optimal allocation. Simulation results are also reported.

D-Optimal Minimax Response Surface Designs

❖ Julie Zhou, University of Victoria, Dept of Mathematics and Statistics, Victoria, BC V8W 3P4 Canada, *jzhou@math.uvic.ca*

Key Words: optimal designs, response surface designs, robust designs, discrete design space

Quadratic functions can provide good approximation to many response functions in small regions of interest. Thus quadratic models are often assumed and used in response surface designs. One classical design criterion is to minimize the determinant of the variance of the regression estimator, and the designs are called D-optimal designs. The bias of the estimator is ignored. To reflect the nature that quadratic models are only approximately true, we propose a robust design criterion to study response surface designs. Both the variance and the bias are considered in the criterion. In particular, D-optimal minimax designs are investigated and constructed. Examples are given to compare D-optimal minimax designs with classical D-optimal designs.

Some Results on Two-Level Nonregular Designs Constructed from Quaternary Codes

❖ Kin Hing (Frederick) Phoa, University of California, Los Angeles, Department of Statistics, Los Angeles, CA 90095-1554, *fredphoa@stat.ucla.edu*; Hongquan Xu, University of California, Los Angeles

Key Words: Nonregular Design, Quaternary Linear Code, Generalized Minimum Aberration, Generalized Resolution, Aliasing Indexes

The research of developing a general methodology for the construction of optimal nonregular designs has been very active in the last 10 years. Recent research by Xu and Wong (2006) suggests a new construction method through quaternary linear codes. In this talk, we explore some properties and uses of quaternary codes towards the construction of nonregular designs. From these properties, we find some theoretic results and derive some applicable formulas to generate the optimal 2^2(n-1) and 2^2(n-2) designs, with respect to the generalized minimum aberration and generalized maximum resolution criteria. The generalization to 2^2(n-k) designs is suggested and some illustrative examples on the applications of these results are presented.

Optimal Designs for Dose-Finding Experiments in Toxicity Studies

❖ Weng-Kee Wong, University of California, Los Angeles, Department of Biostatistics UCLA, 10833 Le Conte AVe, Los Angeles, CA 90095, *wkwong@ucla.edu*; Holger Dette, Ruhr-Universitaet Bochum; Andrey Preplyshev, St. Petersburg State University

Key Words: continuous design, design efficiency, Elfving theorem, local c-optimal design, robust optimal design

We use optimal approximate design theory to construct experimental designs for toxicology studies on a prototype dose interval. The specific goals are to estimate the effective dose of fetal malformation rate, prenatal death rate and the overall toxicity using information on the number of implants, number of prenatal deaths and number of live fetuses. We assume the intra-litter correlation coefficient to depend on the dose and use Weibull models to represent probabilities of malfunction and prenatal death. Locally optimal designs are found, along with analytical conditions on the number of dose levels required for the study. We also compare our designs with popular equidistant designs and investigate sensitivities of our optimal designs to misspecification in the nominal values.

281 Sample Size and Effect Size Estimation ● ✪

Biopharmaceutical Section, Section on Statistical Consulting, WNAR

Tuesday, July 31, 10:30 am–12:20 pm

Sample Size for Equivalence Trials: A Case Study from a Vaccines Lot Consistency Trial

❖ Jitendra Ganju, Amgen Inc., CA 94109, *jganju@biostatworks.com*

Key Words: sample size, power, variance components, vaccines

For some trials simple but subtle assumptions can have a profound effect on the size of the trial. The design of a vaccine lot consistency trial is a case in point. Standard sample size formulas rely on only one component of variation, namely, the variation within lots. The other component, the variation between lots, is either unintentionally ignored or assumed to be equal to zero. But variation between lots, however small, will be inherent in the manufacturing process. Using data from a published lot consistency trial we demonstrate that when the between-lot variation is only 0.5% of the total variation the relative increase in the sample size is nearly 400% when compared to the assumption that the lots are identical.

Assessing the Accuracy of the WMWodds Method for the Wilcoxon-Mann-Whitney Test

❖John Castelloe, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513, *John.Castelloe@sas.com*; Ralph G. O'Brien, The Cleveland Clinic

Key Words: Wilcoxon two-sample test, Mann-Whitney test, power, effect-size

A simple effect-size parameter, WMWodds, allows us to correctly state the hypotheses being tested by the Wilcoxon-Mann-Whitney test and compute associated point estimates, confidence intervals, p-values, and power probabilities. But the underlying theory is asymptotic (Central Limit Theorem, delta method). How well does it behave with moderate and even very small samples and with discrete (ordinal) distributions with just a few categories? We present an extensive Monte Carlo study of the WMWodds-based p-values, the coverage probabilities for its confidence intervals, and the WMWodds-based power probabilities for statistical planning involving the WMW test.

Power and Sample Size Estimation for Comparing Two Means: Assessment and a New Tool

❖Ken Gerow, University of Wyoming, Dept 3332, 1000 E University Ave, Laramie, WY 82071, *gerow@uwyo.edu*

Key Words: sample size, power, paired

Formulae for sample size calculations for statistical power in the literature (especially textbooks) are often predicated on unrealistic assumptions (e.g., equal variances) or unrealistic, or simply undesirable, designs (two independent samples of the same size). In addition, sample size and power calculations usually involve repeated use of any given formula, as the researcher works through varying inputs (alpha, power choices, design choices, variation/mean relationships, etcetera), making the process of considering sample size and power issues unpleasantly tedious and error-prone. In this presentation, I will show (and subsequently make freely available) a rich and interactive sample size calculator that enables calculations for a wide variety of variations on the two-sample setting.

Localization versus Holism in the Analysis of Neurological Imaging Data: Power Comparisons of Multivariate and Univariate Methods

❖Dawson Hedges, Brigham Young University, 1054 SWKT, Provo, UT 84602, *dawson_hedges@byu.edu*; Suzanne Hendrix, Myriad Pharmaceuticals; Bruce Brown, Brigham Young University

Key Words: brain image data, MANOVA, posttraumatic stress disorder

Clinical studies are often burdened by the constraints of small sample size, due to cost considerations, diagnostic issues, and comorbidity. The resultant loss of power can obscure clinically meaningful information. In a brain imaging (MRI) study, univariate analyses showed no significant differences in volumes of 82 brain areas (in 20 main regions) between five healthy controls and six subjects with posttraumatic stress disorder. Two measures of PTSD, the CAPS and a diagnostic interview, were found to correlate highly (r = .952). Surprisingly, a lens model analysis revealed that of the 20 main regions, 11 accounted for over 35% of the diagnostic accuracy, with one accounting for 95%, and several over 50%. Using this information, an appropriate MANOVA analysis was crafted that showed similar high levels of variance accounted for. A strategy is proposed for dealing with small N data.

Small Sample Properties of Information-Based Monitoring

❖David Sawrie, The University of Alabama at Birmingham, Department of Biostatistics, 128 Heather Lane, Pelham, AL 34124, *dsawrie@bellsouth.net*; Christopher S. Coffey, The University of Alabama at Birmingham

In the case of normally distributed outcome measures, sample size depends upon the variance, a nuisance parameter. If variance were known, we could easily plan to collect the exact number of observations required to perform a two sample test at the desired level of power. Unfortunately, true variance is unknown. Recent literature suggests the use of information based monitoring to update an initial variance estimate based upon accumulating study data. Specifically, the literature proposes the integration of information based monitoring into a group sequential framework within which data are already monitored periodically for early indications of efficacy. The approach successfully assures sufficient power and maintains approximately nominal alpha levels for large samples. This paper investigates the effect of this method upon power and type I error when small sample sizes are considered.

282 Atypical Experiments in Industry ● ✪

Section on Physical and Engineering Sciences

Tuesday, July 31, 10:30 am–12:20 pm

Consumer Product Optimization Through a Mixture-of-Mixtures Experiment

❖William Brenneman, Procter & Gamble, 8700 Mason Montgomery Rd, Mason, OH 45040, *brenneman.wa@pg.com*

Key Words: mixture-of-mixture experiment, multiple response approach, desirabitlity function

Mixture designs are used extensively during the formulation stage of many consumer products. There are situations where one or more of the mixture components are made up of sub-components that need to be investigated as well. This leads to what are commonly termed mixture-of-mixtures (MoM) experiments. A real example from industry is presented along with the resulting analysis, complicated by the fact that there are multiple continuous and binary responses.

What Textbook Experiments?

❖Winson Taam, The Boeing Company, POBox 3707 MC 7L22, Seattle, WA 98124-2207, *winson.taam@boeing.com*

Key Words: experimental design, restricted randomization, quantal response, spatial

Engineering problems and investigations in aerospace industry are quite diverse and complex. This poster would summarize several experiments where traditional approaches need to be adjusted to fit industrial settings, or new approaches need to be developed. The cost of industrial experiments varies but always strictly monitored. Some experiments could be executed sequentially and some are confined to feasible subsets. These few experiments shed some light on the commonalities and differences faced in aerospace industry.

Estimating the Probability of Detecting Cracks in Aerospace Structures Using Nondestructive Inspections

❖Peter Hovey, University of Dayton, 300 College Park Dr, Dayton, OH 45415, *peter.hovey@notes.udayton.edu*

Key Words: estimation, fatigue, aerospace, NDI

Nondestructive Inspections are used extensively in the Aerospace industry to look for cracks in structures that would not be readily apparent from a visual inspection. Future maintenance is planned based on the sizes of cracks that may be missed during an inspection. The probability of detecting (POD) cracks as a function of crack size is used as the gauge of the likelihood of a large crack remaining in the structure. Estimating the POD function poses a unique statistical challenge because there are many environmental factors that impact the POD that cannot be duplicated in a planned study designed to estimate the POD function. Several approaches that have been used in the past will be presented and some of the current efforts to provide better estimates will be described.

Design on Nonconvex Regions: Optimal Experiments for Spatial Process Prediction with Applications to Industrial Processes

❖Matthew Pratola, Simon Fraser University, Dept of Statistics, 8888 University Drive, Burnaby, BC V5A 1S6 Canada, *mtpratol@cs.sfu.ca*

Key Words: optimal design, non-convex, spatial, Gaussian, process, ISOMAP

Modeling a response over a non-convex design region is a common problem in many areas of industrial research, such as engineering and geophysics. Unfortunately, the tools available to model and design for such responses are limited as standard methods are not appropriate. Some success has been found by applying the Gaussian Process (GP) model with a non-Euclidean distance metric. A difficulty is that transformation of this metric is required to be able to model a GP over such regions. The questions of how to make this transformation, select design points and fit GP models have received little attention. Our work builds on existing results to propose a valid transformation. A new method for selecting design points with the GP model over non-convex regions is proposed. Optimal designs for prediction are described, and a simulation study demonstrates the improvements that are realized.

Dynamic Structural Health Monitoring Validation

❖David Banaszak, Air Force Research Laboratory, AFRL/VASV Bldg 65, 2790 D Street, Wright Patterson, OH 45344, *david.banaszak@wpafb.af.mil*

Key Words: Calibration, ANOVA, Vibration, Remote Control, Structures, Experiments

The Air Force invented a calibration technique that allows one person to perform multiple, mechanical end-to-end calibrations of structural dynamics measurement systems. This new technique stimulates dynamic measurement transducers contained in structures, with a measurable mechanical input level such as acceleration or stain. The Air Force evaluated the invention in the lab using a commercial off-the-shelf vibration paging system consisting of a master control unit and 8 individual exciters (pagers). The exciters simulate accelerometers mounted on structures. Air Force engineers validated the technique using a completely randomized block design experiment consisting of three 2-level factors: material, structural thickness and excitation mode. The exciter serial numbers are a blocking factor. The Air Force measured amplitude, frequency and transfer function between the accelerometers.

Complex Experimental Design and Simple Data Analysis

❖Joseph G. Pigeon, Villanova University, Department of Mathematical Sciences, 800 Lancaster Avenue, Villanova, PA 19085, *joseph.pigeon@villanova.edu*

Key Words: experimental design, fractional factorial, strip plot

Very often in industrial factorial experiments, not all factors have the same error variance associated with them. This generally arises because it may be more costly or simply not feasible to completely randomize the order of the experimental runs. The result of this restricted randomization is that multiple error measures are introduced into the experiment and the resulting analysis needs to account for these multiple error measures. Experimental designs that combine fractional factorials may have a multiple error structure that is even more complicated than the usual split plot experiment. These designs are sometimes called strip plot or multiway split unit designs. This poster uses a pharmaceutical example to emphasize the need for experimenters to recognize these types of experiments and suggests the use of simple data analysis methods for analyzing these types of experiments.

J-FACT: A Factor-Covering Design Tool for Deriving Software Test Cases

❖Joseph Morgan, SAS Institute Inc., 100 SAS Campus Drive, Cary, NC 27513, *joseph.morgan@sas.com*; Bradley Jones, SAS Institute Inc.

Key Words: Factor Covering Design, Covering Array, Transversal Cover, Design of Experiments, Software Testing

Over the past 20 years, factor covering designs, also known as covering arrays or transversal covers, have been increasingly advocated as a mechanism to derive suites of test cases for testing software products. Typically, the software product is viewed as a k parameter function with the input space of each parameter partitioned into distinct classes (i.e., levels). The challenge is to construct a design, with the fewest number of runs (i.e., test cases), so that the combined input space is effectively exercised. If the design has the property that for any t (t< =k) parameters, all level combinations exist, then the design is referred to as a t-covering design. A t-covering design is "optimal" if the number of runs is minimal for t. We describe how "optimal" t-covering designs may be constructed and illustrate the use of such designs for testing software products.

283 Recent Advances in Microarray Data Analysis III ● ✪

Biometrics Section

Tuesday, July 31, 10:30 am–12:20 pm

ROC Curve of a Generalized Odds-Rate Model

❖Huining Kang, University of New Mexico, HSC DOIM MSC 10 5550, 1 University of New Mexico, Albuquerque, NM 87131, *hukang@salud.unm.edu*; Edward J. Bedrick, University of New Mexico; Susan R. Atlas, University of New Mexico; Cheryl L. Willman, University of New Mexico

Key Words: ROC curve, Proportional odds-rate model, Pseudo maximum likelihood estimator

Many approaches to estimating a receiver operating characteristic (ROC) curve assume a binormal form for it. However, one caveat is that the binormal ROC curve is not concave when the slope parameter is not equal to one, which is unlikely for most practical problems. We propose a new functional form for the ROC curve derived from the generalized proportional odds-rate (GPO) model. While sharing some properties with the binormal ROC curve, the GPO ROC curve has an advantage of being concave everywhere. One of the parameters can be interpreted as the odds ratio of diseased versus non-diseased subjects, which makes the GPO ROC curve useful in analyzing the effect of covariates on diagnostic tests. We have developed a pseudo maximum likelihood estimator for estimating the GPO ROC curve and illustrate the method with a data set from a gene expression microarray study on acute leukemia.

Transformation for the Additive Effect on the Affymetrix Probe Data

❖ Fenghai Duan, University of Nebraska Medical Center, 984350 Nebraska Medical Center, Omaha, NE 68198-4350, *fduan@unmc.edu*

Key Words: Affymetrix, Transformation, probe

The microarray technology enables us to measure the expressions of thousands of transcripts in a single experiment. The raw intensity value of a gene is the measurement of the abundance of the corresponding mRNA. The logarithmic transformation is typically performed to the raw intensity data before any further analysis. Some researches have been conducted to search for other transformation functions, mostly focusing on the variance stabilization. In this study, we were using an Affymetrix dataset to investigate the issue of additivity, which is particular important in summarizing the probe level data. A statistical model was also proposed to accommodate our findings. Our results showed our approaches were useful in the quality assessment, normalization and summarization at the probe level for Affymetrix data.

Variable Selection in Penalized Model-Based Clustering with Cluster-Dependent Diagonal Covariances with Application to Microarray Data

❖ Benhuai Xie, The University of Minnesota, A460 Mayo Building MMC 303, 420 Delaware St SE, Minneapolis, MN 55455, *xiex0041@umn.edu*; Wei Pan, The University of Minnesota; Xiaotong Shen, The University of Minnesota

Key Words: Penalization, Mixture model, EM algorithm, BIC

Recently Pan and Shen proposed a penalized model-based method to realize variable selection in cluster analysis for high-dimensional data. However, their method requires a common diagonal covariance matrix across clusters. Here we consider a more general situation with cluster-dependent diagonal covariance matrices, for which both mean and variance parameters have to be penalized to realize variable selection. An EM algorithm is derived to obtain the maximum penalized likelihood estimates. Numerical examples are provided to demonstrate the utility of the proposed method.

FDR Control in Identification of Differentially Expressed Genes Using Mixed Linear Models When Some Variance Components May Be Zero

❖ Cumhur Demirkale, Iowa State University, Department of Statistics, Snedecor Hall, Ames, IA 50011-0001, *cyusuf@iastate.edu*; Tapabrata Maiti, Iowa State University; Dan Nettleton, Iowa State University

Key Words: Microarray data analysis, False Discovery Rate, Mixed linear models, Variance components

In a microarray experiment, one experimental design is used to obtain expression measures for all genes. One popular analysis method involves fitting the same mixed linear model for each gene, obtaining gene-specific p-values for tests of interest involving fixed effects, and then choosing a threshold for significance that is intended to control False Discovery Rate (FDR) at a desired level. When some random factors have zero variance components for some genes, the standard practice of fitting the same full mixed linear model for all genes can result in failure to control FDR. We propose a new method which combines results from the fit of full and selected mixed linear models to identify differentially expressed genes and provide FDR control at target levels when the true underlying random effects structure varies across genes.

Controlling the Average False Discovery in Large-Scale Multiple Testing

❖ Meng Du, University of Toronto, Toronto, ON Canada, *meng@utstat.utoronto.ca*; Muni S. Srivastava, University of Toronto

Key Words: average false discovery, DNA microarrays, false discovery rate, multiple testing, powers, p-values

One of the most important issues in large-scale multiple testing is how to choose an appropriate false positive rate. In this paper, we define the average false discovery (AFD) as the expected proportion of false positives among all of the true cases. The appropriateness of the AFD is addressed in comparison with other false positive rates include the false discovery rate (FDR). We also propose an AFD controlling procedure. The theoretical power of the AFD procedure is derived under reasonable assumptions. The performance of the AFD procedure is compared to that of the Benjamini and Hochberg (1995) FDR controlling procedure through simulation. The simulation results show that the AFD procedure is superior to the Benjamini and Hochberg procedure in the sense of having more stable powers irrespective of the proportion of the true cases.

Integrated Clustering of Heterogeneous Genomic Data

❖ Jyotsna Kasturi, Johnson & Johnson PRD, 1000 Route 202 South PO Box 300, Raritan, NJ 08869, *jkasturi@prdus.jnj.com*

Key Words: clustering, heterogeneous data, information fusion, gene expression, motif frequency

The availability of high throughput technologies facilitates the inference of gene regulatory mechanisms on a global scale via genome-wide studies. Data from a single microarray study alone, though information-rich, lacks the specificity to identify these complex mechanisms of action, and often need to be combined with data from multiple microarray studies as well as other genomic sources such as DNA binding motifs and gene ontologies. Analytical methods that perform integrated analysis on such data are much needed. An unsupervised clustering algorithm to perform information fusion from heterogeneous genomic data to identify clusters of genes using the combined data is introduced. A weighting scheme allows the influence of each data type on the final clustering to be specified a priori. Results on real data show that this approach identifies co-regulated genes effectively.

Reconstruction of Gene Transcription Regulation Network Using a Bayesian Approach

❖ Feng Hong, University of Illinois at Urbana-Champaign, 101 Illini Hall MC 374, 725 S Wright St, Champaign, 61820, *fenghong@uiuc.edu*

Key Words: microarray, MCMC, bioinformatics

Gene transcription regulation is a fundamental biological process and understanding the regulatory relationship between the proteins and their regulated genes is a center problem in biology. High-throughput data such as microarray gene expression measurement, protein-DNA binding data under various conditions such as gene disturbance now become available to help reconstructing the regulation network. In this paper we use Bayesian modeling to integrate the information from these different data source. Markov chain Monte Carlo is used to estimate parameters in the model. We will demonstrate our method in simulated data and real data.

284 Design and Sampling Methods ● ✪

Biometrics Section

Tuesday, July 31, 10:30 am–12:20 pm

Optimal Designs in High-Throughput Screening

❖ Xianggui Qu, Oakland University, 2200 N Squirrel Road, Rochester, MI 48309, *qu@oakland.edu*

Key Words: Binary row-column designs, high throughput screening, (m,s)-optimality, row-column designs

A new class of row-column designs is proposed for high throughput screening. These designs are minimally treatment-connected row-column designs. That is, all treatment contrasts are estimable regardless of the row and column effects. The (m,s)-criterion is used to select optimal designs. (M,s)-optimal designs are constructed over a class of square binary designs. Nonsquare optimal designs and other alternatives are also discussed.

Sampling Design for a Survey of Pregnant Women Following a Natural Disaster

❖ Jason Hsia, Centers for Disease Control and Prevention, 4770 Buford Hwy NE, MS K21, Atlanta, GA 30341, *jhsia@cdc.gov*; Marianne E. Zotti, Centers for Disease Control and Prevention; Van T. Tong, Centers for Disease Control and Prevention

Key Words: Cluster Sampling, Natural Disaster, Response Rule

In pregnancy-related surveys, we often find them through a screening of reproductive-age women or from newborn records. However, if the study period is following a natural disaster, this approach may not be practical or possible. Because < 5% of reproductive-age women are pregnant and a natural disaster is often short, an extraordinarily large number of reproductive-age women would need to be screened to identify enough pregnant women. During a natural disaster, residents are often more mobile, so it may not be feasible to locate pregnant women from birth certificate information. To overcome this difficulty, we propose a modified multistage cluster sampling. We start with a conventional multistage cluster sample. At the last stage, we will let sampled women report all pregnant women they know within their cluster. We will then interview all pregnant women being reported.

Adaptive Two-Stage Designs for Heterogeneous Patient Populations in Phase II Cancer Trials

❖ Karen Messer, University of California, San Diego, Division of Biostatistics, 3855 Health Sciences Dr MC 0901, La Jolla, CA 92093-0901, *kmesser@ucsd.edu*; Michael Donohue, University of California, San Diego

Key Words: Phase II trials, group sequential design, two stage design, adaptive design

In a Phase II cancer trial there may be several patient groups. For example, in a Leukemia trial patients in 1st relapse may have 30% probability of response under standard treatment while patients in 2nd + relapse may have only 10% probability of response. In many cancer trials the experimental agent is expected to induce severe toxicity. Such trials are usually designed with a stopping rule which checks for initial evidence of efficacy after a first stage enrollment target is met. If there is insufficient evidence, the trial stops early for futility. We discuss the standard two-stage designs in this situation, and describe their operating characteristics under heterogeneous patient enrollment. Simple, approximately optimal designs which account for heterogeneity are presented. We recommend a practical adaptive design strategy which we have implemented at Moores UCSD Cancer Center.

Inverse Probability-Weighted Estimators in the Stratified Nested Case-Control Sampling Methods

❖ Sachiko Tanaka, Tokyo University of Science, Faculty of Engineering Course, 1-14-6, Kudan-kita, Chiyoda-ku, Tokyo, 102-0073 Japan, *sachiko@ms.kagu.tus.ac.jp*; Yutaka Matsuyama, University of Tokyo; Shiro Tanaka, University of Tokyo; Seiichiro Yamamoto, National Cancer Center

Key Words: absolute risk, counter-matching, matching, nested case-control study, weighted estimators

Stratified nested case-control designs are widely used in the expectation that these methods may increase precision relative to simple nested case-control design. Although conditional logistic regression is used for analysis of data with matching, it will be inefficient because it does not use all subjects at risk. Furthermore, usual method does not give the estimate of absolute risk. In this talk, we introduce the inverse probability weighted estimator of Cox regression parameter under stratified sampling designs. We estimate selection probability and conditional expectation of missing covariate using nonparametric method that incorporates information of matching factors. Estimates of absolute risk are obtained based on Nelson-Aalen estimators. We also extend our method to counter-matching design. The performance of the proposed method is investigated through simulation studies.

Precision of Double Sampling with Error-Prone and Error-Free Measures and Optimum Case-Control Study Design

❖ Jun-mo Nam, National Cancer Institute, 6120 Executive Blvd, EPS / Room 8028, Rockville, MD 20892, *namj@mail.nih.gov*

Key Words: Double sampling, Efficiency, Fallible classification, Gold standard, Optimum sample size determination, Validation sub-study

We consider double sampling in which cases and controls are classified as exposed or not-exposed by an inexpensive but error-prone measurement in the primary sampling and a random sub-sampling is classified additionally using the expensive error-free measurement. We present the optimum sample sizes which maximize precision of a true odds ratio estimator for a fixed total cost and also those sample sizes minimizing a total cost for a given value of precision. Examples are given using a sudden infant death syndrome and maternal use of antibiotics study and an asbestos-exposure and lung cancer study. A gain in precision and a cost-saving resulted from the optimum design are substantial when a relative cost of the gold standard is high. The optimum double sampling is also compared with a single sampling where subjects are classified by the gold standard exclusively for a fixed total cost.

Early Termination of Two-Stage, Phase II, One-Sample Cancer Clinical Trials

❖ William E. Barlow, Cancer Research and Biostatistics, 1730 Minor Ave, Suite 1900, Seattle, WA 98101, *williamb@crab.org*

Key Words: Phase II, clinical trial, early termination, negative binomial

Two-stage Phase II single sample designs (Simon, 1989) can result in early termination if response is less than expected. If the trial proceeds to Stage 2, the MLE of the response rate is biased , but a UMVUE estimator can be constructed (Jung & Kim, 2004). We propose a fixed sample size for the first stage, but a variable sample size for the second stage. The trial terminates at the maximum allowed sample size or when the number of responses results in rejection of the null hypothesis in favor of the alternative. We use a modification of the Jung & Kim estimator that allows for a truncated negative binomial model in the second stage resulting in an unbiased estimator. Exact confidence intervals may be computed. This design has an

expected sample size smaller or equal to that of the fixed design, but the same power and test size.

Design of Long-Term Viral Dynamic Studies in Semiparametric Mixed-Effects Models

❖Jeong-Gun Park, Harvard School of Public Heath/Frontier Science & Technology Research Foundation, Boston, MA 02115, *park@sdac.harvard.edu*; Yangxin Huang, University of South Florida; Yiliang Zhu, University of South Florida

Key Words: HIV dynamics, Nonlinear mixed-effects Models, Long-term viral dynamics

Studies of HIV dynamics in AIDS research are very important for understanding the pathogenesis of HIV infection and for assessing the potency of antiviral therapies. Nonlinear parametric models, derived from the mechanisms of HIV infection and drug action, have been used to fit short-term clinical data from AIDS clinical trials, but they are not adequate to fit the long-term viral dynamic data. Here we use a simulation-based approach to deal with design problems for long-term longitudinal HIV dynamic studies under the framework of semiparametric nonlinear mixed-effects models which can not only preserve the meaningful interpretation of the short-term HIV dynamics, but can also characterize the long-term viral dynamics. We compare a finite number of feasible candidate designs numerically and provide guidance on how a design might be chosen in practice.

285 The Era of Omics Data: Statistical Issues and Possible Solutions ● ✪

Biopharmaceutical Section, Biometrics Section

Tuesday, July 31, 10:30 am–12:20 pm

Proteomic Analysis by 2-D PAGE for Response to Low-Dose Ionizing Radiation and Arsenic Exposure

❖Dan Li, University of California, Davis, 6600 Orchard Park Circle, Apt 6024, Institute for Data Analysis and Visualization, Davis, CA 95616, *dali@ucdavis.edu*; Susanne R. Berglund, University of California, Davis; Alison Santana, University of California, Davis; Zelanna Goldberg, University of California, Davis; David Rocke, University of California, Davis

Key Words: Protein expression data, 2-D PAGE, Statistical differential analysis, Empirical Bayes estimate, Missing values, EM algorithm

Two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) is an important technique in proteomics. We have developed a procedure to analyze protein expression data generated by 2-D PAGE to identify the proteins that are changed between biological samples taken under different conditions. The procedure consists of data normalization, variance stabilization, analysis of variance, estimate of protein-specific variance, adjustment for multiple comparisons and a Chi-square test. Missing data is a common problem encountered by researchers when using 2-D PAGE. Each missing spot is either left-censored or just randomly missing. The application of EM algorithm to statistical differential analysis with such missing data is studied. A case study on the effect of combined exposures of low dose arsenic and low dose ionizing radiation is presented.

Statistical Differential Analysis of 2-D Gel Images

❖Feng Li, University of Maryland, Baltimore County, 4702 Gateway Terrace, Apt C, Baltimore, MD 21227, *usalf2002@yahoo.com*; Francoise Seillier-Moiseiwitsch, Georgetown University Medical Center

Key Words: 2-D Gel, image analysis, multiple comparison, FDR, empirical Bayes

s2-D Gel is the most powerful protein separation technique. Statistical differential analysis of 2-D Gel images is important and necessary but it is in its infancy. To do statistical differential analysis of 2-D Gel images, the following procedure is used. The 2-D Gel images must be well aligned first. Watershed algorithm was applied to selected 2-D gel images to build a common watershed region for all images. Each watershed region was classified into background and object parts. The difference between the means of the object part and background part was used as a summary statistics for each watershed region. ANOVA method was applied to each watershed region to get a p-value. To account for multiple hypothesis testing problem, FDR procedure and empirical Bayes method were used to detect meaningful results. The application of this procedure to a set of gel images showed the effectiveness.

Methods for Classification of NMR Metobololomic Spectroscopy

❖Yuanxin Xi, University of California, Davis, Institute of Data Analysis and Visualization, 5000 Orchard Park Circle, Apt 5811, Davis, CA 95616, *xi11west@gmail.com*; David Rocke, University of California, Davis

Key Words: NMR, Adaptive Binning, Classification, Biomarker Discovery

We present a classification case of 2D-NMR TOCSY datasets for biomarker discovery in the study of Pseudoxanthoma elasticum (PXE), a heritable recessive connective tissue disorder. PXE is primarily caused by metabolic disorder and we use in-vivo 2D-TOCSY spectroscopy to detect small molecular metabolites for 21 patients and 19 controls. The datasets are preprocessed by an adaptive binning based on peak probability densities and then compared by logistic regression and support vector machines. The results show this adaptive binning method is more stable than classical fixed size binning. The two classes are well separated in PLS component spaces and the loading plot indicates the significant regions that suggest potential biomarkers.

Genetic Association Analysis Using the Penalized Likelihood Approach

❖Xiwu Lin, GlaxoSmithKline, 1250 S Collegeville Rd, Collegeville, PA 19426, *xiwu_2_lin@gsk.com*; Kijoung Song, GlaxoSmithKline; Daniel Parks, GlaxoSmithKline; Jie Cheng, GlaxoSmithKline; Kwan Lee, GlaxoSmithKline

Key Words: penalized likelihood, genetic association, penalty function, SNP

With the advances of genotyping technologies, genome-wide association studies with hundreds of thousands of single nucleotide polymorphisms (SNPs) are available. To analyze such kind of genetic data, one usually analyzes each SNP one at a time to examine the association between each SNP and disease status. One big challenge for single SNP approach is the issue of multiple testing. Another issue for such approach is that multiple SNPs might have separately small effects but jointly they might have a large effect on the disease of interest. To overcome the above issues, we use a penalized likelihood approach to analyze the effects of all SNPs on a chromosome level. We propose a weighted penalty function to incorporate the relationship among SNPs. The performance of the proposed method is evaluated using simulation data. An example based on a genetic epidemiology data will be provided.

Microarray Experiments for Pharmacogenomics

❖Youlan Rao, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, *rao@stat.ohio-state.edu*; Yoonkyung Lee, The Ohio State University; Jane Chang, Bowling Green State University; Tao Wang,

University of South Florida; Jason C. Hsu, The Ohio State University

Key Words: Microarrays, prognostics, mixed linear model, multiple testing, classification, sample size

The development of a microarray for clinical prognosis is envisioned as a two-stage process. The first stage is to find marker genes and train a prognostic algorithm based on data with known disease outcomes. The second stage is to validation the prediction algorithm using arrays with marker genes only. In this talk, we show that if a microarray experiment is designed according to statistical principles to avoid bias and allow the variance components of subject, sample, and noise to be estimated separately, then sample sizes of subject, samples within subject, and replication of probes can conceptually be computed to meet prespecified sensitivity and specificity requirements of the device.

Analyzing Omics Data Using rSVD and NMF

❖ Li Liu, sanofi-aventis, 1041 Rt 202 206, Bridgewater, NJ 08807, *li.liu@sanofi-aventis.com*; Donghui Zhang, sanofi-aventis

Key Words: rSVD, NMF, omics data

Robust singular value decomposition (rSVD) and non-negative matrix factorization (NMF) are two useful methods for analyzing omics data. We compared the two methods using simulated data and real data. The similarities and differences between the two methods are studied, and some practical remarks are provided.

A New Procedure of Identifying Differentially Expressed Genes in Microarray Experiments

❖ Chunrong Dong, Case Western Reserve University, 2013 Random Rd, Down, Cleveland, OH 44106, *cxd55@cwru.edu*

Key Words: microarray, gene expression, hypothesis test

Based on experimental microarray data, a novel procedure was proposed for identifying differentially expressed genes under independence, heterogeneous variance hypothesis of linear models.

286 Bayesian Methods

ENAR, Biometrics Section

Tuesday, July 31, 10:30 am–12:20 pm

Comparing Various Designs of Temporal Microarray Experiments

❖ Yijun Yang, Boston University, 16 Walnut St Apt 42, Arlington, MA 02476, *yijunyang02190@yahoo.com*; Paola Sebastiani, Boston University

Key Words: time series, gene expression, clustering

Common temporal gene expression studies consist of a number of unevenly spaced observations that are often replicated to average out biological and sampling variability. Due to the high cost of microarray experiments, duplication of the same time point experiments are often chosen at the price of the sampling frequency. The sampling frequency is important to capture the essential dynamics of the data generation process and to increase the power to identify interesting genes with similar expression profiles. We conducted a simulation study to investigate the effect of replicating microarray experiments compared to increasing the sampling frequency, and we determined the minimum number of time points for recovering true number of clusters. CAGED (cluster analysis of gene expression dynamics) was used for clustering the simulated gene expression data.

Bayesian Analysis of Cross-Classified Spatial Data with Autocorrelation

❖ Xiaolei Li, GlaxoSmithKline, 812 Morgan Dr, Royeersford, PA 19468, *xiaolei06@gmail.com*; Murray K. Clayton, University of Wisconsin-Madison

Key Words: MCMC, Spatial Statistics, Categorical Data, Gibbs Sampler, Cross-classification, high-dimentional parameter space

The work is focused on the development and application of statistical methodologies to the analyses of categorical data collected over space. When several spatial attributes are considered simultaneously, their mutual associations are hard to characterize. The standard chi-squared analysis becomes invalid and can lead to wrong conclusions because of the spatial autocorrelation within each attribute. Our methods focus on identifying the mutual independence between two multi-categorical spatial processes over a finite lattice. Multinomial autologistic Markov models are constructed for more than one multi-categorical spatial processes as well as the mutual dependence between any two of them. For model inferences, a new MCMC algorithm are proposed for estimations in high-dimensional parameter space and combined with Gibbs sampler. Then, this Bayesian procedure is justified theoretically.

Bayesian Sample Size Determination Under Hypothesis Tests

❖ Gary Cutter, The University of Alabama at Birmingham, 1665 University BLVD Suite 414, Birmingham, AL 35294, *cutterg@prodigy.net*; Xiao Zhang, The University of Alabama at Birmingham

Key Words: Bayes factor, Hypothesis test;, Conjugate distribution, Simulation, Sensitivity

We develop a Bayesian approach for calculating sample sizes for clinical trials using the framework of hypothesis tests. The hypotheses we consider are generic and involve specifying conjugate distributions for treatment effects or group difference, and variances of the data. We select sample sizes using the Bayes factor and the averaged type I error and type II error defined by Weiss [1]. Utilizing conjugate distributions, the hypothesis allows the uncertainty inherent in both treatment effects and variances in the estimates from the data. Therefore our approach for calculating sample sizes permits informative prior information for unknown quantities through the hypothesis specification. We investigate the sensitivity of calculated sample sizes to the prior information through simulation. The illustrated example is from an MS clinical trial and the computation is done through WINBUGS1.4

Bayesian SUR Modeling of Multiple Ordinal Traits for Genome-Wide Epistatic QTL Mapping

❖ Samprit Banerjee, The University of Alabama at Birmingham, Dept of Biostatistics, Section on Statistical Genetics, 1665 University Blvd, Ryals Public Health Bldg, Room 327, Birmingham, AL 35294, *samban@uab.edu*; Nengjun Yi, The University of Alabama at Birmingham

Key Words: Seemingly Unrelated Regression, Bayesian Model Selection, QTL mapping, multiple traits, epistasis

Unraveling the genetic etiology of complex traits by identifying complex epistatic quantitative trait loci (QTL) across the entire genome, still poses a bewildering challenge to contemporary statistical geneticists. The majority of the existing methods focus on a single trait, even though typically data on more than one phenotype are collected. In this paper, we present a multiple trait Bayesian composite model space approach to perform Seemingly Unrelated Regression (SUR) on complex traits in experimental crosses from two inbred lines. The joint analysis has several advantages over single trait analysis, including the expected improvement in statistical power to detect

QTL and in precision of parameter estimation. It can also provide us with a greater insight in the nature of genetic correlations in certain regions of the genome by testing plieotropy and plieotropy vs. close linkage.

Bayesian Inferences for Bivariate Positive Stable Frailty

❖Madhuja Mallick, Merck & Co., Inc., 126 E Lincoln Avenue, Rahway, NJ 07065, *madhujamallick@hotmail.com*; Nalini Ravishanker, University of Connecticut; Nandini Kannan, University of Texas at San Antonio

Key Words: dependence, baseline hazard

The univariate shared frailty models have several drawbacks. In many applications, the situation arises where the subjects in the same group possess different frailty random components rather than sharing similar frailty components within the group. The notion of multivariate frailty enables us to overcome the difficulties of univariate shared frailty models. This presentation proposes bivariate positive stable frailty model to incorporate more heterogeneity than the shared positive stable frailty model. The estimation procedure for bivariate positive stable density is not easy due to lack of closed form of the bivariate positive stable density function. The Bayesian inference for the bivariate positive stable frailty model for multivariate survival data is described here.

Bayesian Analysis of the Effect of Intentional Weight Loss on Mortality Rate

Nengjun Yi, The University of Alabama at Birmingham; Shouluan Ding, The University of Alabama at Birmingham; ❖Scott W. Keith, The University of Alabama at Birmingham, Department of Biostatistics, 1665 University Blvd, Birmingham, AL 35294, *swkeith@uab.edu*; Christopher S. Coffey, The University of Alabama at Birmingham; David B. Allison, The University of Alabama at Birmingham

Key Words: Bayesian analysis, latent variables, mortality, obesity, weight loss

The effect of weight loss (WL) on mortality rate is widely studied in obesity research. Separating the effects of intentional weight loss (IWL) from unintentional weight loss (UWL) continues to be a challenge. It has been shown that WL among people intending to lose weight is not equivalent to IWL. We constructed Bayesian latent variable linear models that allow the separation of IWL and UWL effects among those intending to lose weight by augmenting their unobserved UWL with information from observed WL among those not intending to lose weight. This approach provides estimates of IWL and UWL effects as well as any other parameters of interest. We applied our method to a real rodent caloric restriction study dataset. Our results suggest that IWL has a substantial beneficial effect on mouse lifespan, in contrast to UWL. We also discuss extensions to human data and censored outcomes.

Testing Random Effects in the Linear Mixed Model Using Bayes Factors

❖Benjamin Saville, The University of North Carolina at Chapel Hill, 3106 McGavran Greenberg, CB 7420, Chapel Hill, NC 27599, *bsaville@bios.unc.edu*; Amy H. Herring, The University of North Carolina at Chapel Hill

Key Words: random effects, linear mixed model, Bayes factors

Deciding which predictor effects may vary across subjects is a difficult modeling issue. Testing on the boundary of the parameter space changes the asymptotic distribution of some classical test statistics and causes numerical problems for Bayesian methods of approximating Bayes factors. Random effects also induce high-dimensionality, which can limit the performance of popular approximations to Bayes factors. We propose a simple approach for testing random effects in the linear mixed model using Bayes factors. We introduce factor loadings on the random effects and scale the random effects to the residual variance. We suggest default priors on the factor loadings, and integrate out the random effects and variance components using closed form solutions. We use Laplace's method to approximate the marginal likelihoods needed to evaluate the Bayes factor.

287 Bayesian Model Selection, FDR, and HIV Applications ● ✪

Section on Bayesian Statistical Science

Tuesday, July 31, 10:30 am–12:20 pm

Bayesian Model Selection in High-Dimensional Genetic Association Studies

❖Melanie Wilson, Duke University, Box 90251, Durham, NC 27708, *melanie.wilson@duke.edu*

Key Words: Genetic Association Studies, Bayesian Model Selection, Hierarchical Models

Modern genotyping techniques allow vast amounts of data to be collected for genetic association studies. With this increase in data comes an increased need for statistical models that are able to sort through a large number of correlated covariates as possible disease predictors. To date, the results of such genetic association studies have been disappointing due to the focus on marginal association that over-simplifies the complex etiology of common disease. It is clear that studies of genetic variation and disease must account for the structure and function of genetic pathways. We propose Bayesian hierarchical model selection techniques that search gene-environment and gene-gene interactions in a computationally efficient manor while aiming to strike a balance between model complexity and analytical simplicity. We then compare methods using simulated and real-world datasets.

A Bayesian Multiple Comparison Procedure for Unbalanced Mixed Models

❖Junfeng Shang, Bowling Green State University, 450 Math Sciences Building, Dept of Mathematics and Statistics, Bowling Green, OH 43403, *jshang@bgnet.bgsu.edu*; Farroll T. Wright, University of Missouri-Columbia; Joseph Cavanaugh, The University of Iowa

Key Words: Bayesian method, unbalanced mixed model, imputation, hierarchical model

We propose a Bayesian hierarchical model for multiple comparisons in a mixed modeling framework under a simple-order restriction. We employ Markov Chain Monte Carlo (MCMC) methods to estimate parameters and to obtain estimates of the posterior probabilities that any two of the means are equal, which allows one both to determine if these two means are significantly different and to test the homogeneity of all of the means. In unbalanced data, the behavior of the model is explored along with multiple imputations for values missing at random. Our simulation and application results exhibit that our proposed hierarchical model can effectively unify parameter estimation, tests of hypotheses, multiple imputations for unbalanced data, and multiple comparisons in one setting.

Using FDR To Address a Controversy About the Meaning of P Values

❖Hormuzd Katki, National Institutes of Health, NCI, 6120 Executive Blvd EPS8044, MSC 7244, Rockville, MD 20852-4910, *katkih@mail.nih.gov*

Key Words: Bayes Factor, False Discovery Rate, Local False Discovery Rate, p-value, Evidence, Statistical Significane

Scientists use P values as measures of statistical evidence, but statisticians continue to debate this issue. In particular, two different measures of statistical evidence disagree on the effect of sample size on the meaning of P values, leading to a paradox. These measures have a distinguished history and are related to the measures of evidence implicitly used by the Bayesian versions of False Discovery Rates (FDR) and Local False Discovery Rates (fdr). Considering the False Nondiscovery Rate (FNR) throws light on the paradox as well as an epidemiology challenge put forth to the presenter. The examples bring out the difficulties and irreconcilable differences between FDR/FNR and fdr in how they interpret the evidence represented by P values.

Modeling and Computation for Parametric Bayesian Multiple Hypothesis Testing

❖James Scott, Duke University, ISDS Duke University, Box 90251, Durham, NC 27708, *james@stat.duke.edu*; James Berger, Duke University

Key Words: multiple hypothesis testing, stochastic search, objective priors, variable selection

Multiple hypothesis testing arises in areas as diverse as epidemiology, brain imaging, genomics, and clinical trials, with perhaps the canonical example being DNA microarrays. The standard Bayesian parametric approach is to assume that all hypotheses are independently true or false with a common prior inclusion probability p, which is itself a model parameter to be estimated. This induces shrinkage of all posterior inclusion probabilities toward the posterior mean of p, yielding an automatic penalty for testing more hypotheses. This paper considers extensions of this model to general variable selection problems by relaxing the assumption of independence. We also describe a stochastic search algorithm for exploring the resulting high-dimensional model space, and we recommend prior distributions for model parameters that dovetail nicely with the search strategy.

Nonparametric Bayesian Estimation of Positive False Discovery Rates

❖Anindya Roy, University of Maryland, Baltimore County, Dept of Mathematics and Statistics, Baltimore, MD 21250, *anindya@math.umbc.edu*; Subhashis Ghoshal, North Carolina State University; Yongqiang Tang, SUNY Health Science Center

Key Words: Dirichlet process, Dirichlet mixture, Markov chain Monte Carlo, Multiple testing, Posterior

We propose a Dirichlet process mixture model (DPMM) for the P-value distribution in a multiple testing problem. The DPMM allows us to obtain posterior estimates of quantities such as the proportion of true null hypothesis and the probability of rejection of a single hypothesis. We describe a Markov chain Monte Carlo (MCMC) algorithm for computing the posterior and the posterior estimates. We propose an estimator of the positive false discovery rate based on these posterior estimates and investigate the performance of the proposed estimator via simulation. We also apply our methodology to analyze a leukemia dataset.

A Hierarchical Bayesian Approach in Viral Dynamic System-Based Differential Equation Models with Application to AIDS Studies

❖Yangxin Huang, University of South Florida, Department of Epidemiology and Biostat, College of Public Health MDC56, Tampa, FL 33612, *yhuang@health.usf.edu*; Hulin Wu, University of Rochester

Key Words: Bayesian mixed-effects models, long-term HIV dynamics, longitudinal data, MCMC, time-varying drug efficacy

HIV dynamic studies have significantly contributed to the understanding of HIV pathogenesis and treatment strategies. However, the models of existing studies are mostly developed to quantify short-term dynamics and may not correctly describe long-term virological response due to the difficulty of establishing a relationship of antiviral response with multiple treatment factors. We develop a mechanism-based nonlinear differential equation models with incorporating PK, drug resistance and adherence for characterizing long-term viral dynamics. A Bayesian nonlinear mixed-effects modeling approach is investigated for estimating dynamic parameters by fitting the model to viral load data from an AIDS trial. Some interesting results are presented. These results suggest that dynamic parameters play an important role in understanding HIV pathogenesis, designing new AIDS treatment strategies.

Lipid Analysis for HIV Patients: Does Fasting Bias Matter?

❖Xiting Yang, Carnegie Mellon University, Department of Statistics, Baker Hall 132, Pittsburgh, PA 15213, *xitingy@stat.cmu.edu*; Joseph B. Kadane, Carnegie Mellon University; Heidi M. Crane, University of Washington

Key Words: Bayesian, Hierarchical, Lipid measurement, HIV

This study applies a Bayesian method for investigating the effects of uncertainty of fasting time on lipid measurement for HIV-infected patients and examines whether or not repeated measurement with eight hours fasting is necessary. A hierarchical structure is used to model the effects of fasting, with the posterior distributions calculated using Gibbs sampling and Metropolis-Hasting sampling. The necessity of repeated measurement with mandatory fasting is summarized by checking the distribution of the posterior for the hypothetical fasting lipid measurements for different observed measurement ranges.

288 Sampling Issues and Variance Estimation ●

Section on Government Statistics, Section on Physical and Engineering Sciences, Section on Survey Research Methods

Tuesday, July 31, 10:30 am–12:20 pm

Effects of Calculating Standard Errors Using Sampling versus Variance Weights for 2004 National Nursing Home Survey

❖Rong Cai, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *rcai@cdc.gov*; Sarah Gousen, National Center for Health Statistics; Iris Shimizu, National Center for Health Statistics; Robin Remsburg, National Center for Health Statistics

Key Words: standard errors, sampling weights, variance weights, ratio adjustment, NNHS

The NNHS (National Nursing Home Survey) includes two weights. A sampling weight is used to calculate estimates and a variance weight is used to determine sampling errors of the estimates. Sampling weights are the product of ratio adjustment factors and variance weights. Software restrictions in programs such as SUDAAN, require two separate analyses to get both the estimates and their standard errors, which often confuses analysts and can be a source of error. Therefore, the differences between sampling errors based on the two sets of weights are investigated. The effects of using sampling weights instead of variance weights to compute the sampling errors in analysis are evaluated and corrections are offered to minimize erroneous analytic results.

Variance Estimations for International Price Program Indexes

❖ Te-Ching Chen, Bureau of Labor Statistics, 2 Massachusetts Avenue NE Room 3655, DMI, Washington, DC 20212, *tchen@digitalmanagement.com*; Patrick Bobbitt, Bureau of Labor Statistics; James A. Himelein Jr., Bureau of Labor Statistics ; Steven P. Paben, Bureau of Labor Statistics; Moon Jung Cho, Bureau of Labor Statistics; Lawrence Ernst, Bureau of Labor Statistics

Key Words: Variance estimation, Bootstrap, Jackknife, Balance Repeated Replication (BRR), Certainty Sampling Units, Taylor Series

The International Price Program (IPP) collects data on United States trade with foreign nations and publishes monthly indexes on import and export prices of U.S. merchandise and services. Recently, the IPP evaluated different variance estimation methods such as Taylor Series Linearization, bootstrap, jackknife, and BRR, for their applicability to the IPP. We constructed an artificial universe of monthly price changes for items constructed from 13 years of IPP historical data. We then compared the bias and stability of the variance estimation methods for month-to-month, annual, and long-term price changes by drawing from the universe 1000 samples in various merchandise strata.

Applying M-Estimation to Computation of Price Indexes

❖ Zdenek Patak, Statistics Canada, 100 Tunneys Pasture Driveway, Ottawa, ON K1K4V3 Canada, *zdenek.patak@statcan.ca*

Key Words: Price Index, Robust Estimation, M-estimator

National statistical agencies produce a number of key economic indicators, including the Consumer Price Index and the Service Producer Price Index. These indexes are used by various bodies of government and businesses for setting economic policies, contract negotiations, and the deflation of industrial output. Thousands of observations feed into the index calculations and outlier resistant and efficient methods of estimation need to be developed. In this paper, we propose the use of M-estimators, both in the infinite population and complex survey design settings, to produce timely, efficient, and outlier resistant estimates of price indexes while reducing the amount of manual intervention. A small simulation study compares the behavior of the proposed estimator with a typical fixed weighted index.

Can a Geographic Sort Improve Hot Deck Donor Imputation in the Canadian Census?

❖ Darryl Janes, Statistics Canada, 100 Tunneys Pasture Driveway, RH Coats 15A, Ottawa, ON K1A 0T6 Canada, *darryl.janes@statcan.ca*

Key Words: CANCEIS, Editing, Donor Imputation, Geographic sort

The quality of donor imputation relies upon the content of the chosen donors. The search for donors can be computationally expensive for an imputation system, so it is beneficial to find good donors quickly. A good donor usually has certain characteristics in common with the record requiring imputation. The Canadian Census Edit and Imputation System (CANCEIS) imputes data at the dwelling, family, or person level, and uses a ripple search in stages to find its donors. In the Canadian Census, data is sorted geographically so that nearby households or persons are the first to be considered as donors in hopes of improving imputation quality. This paper will examine the effect of the geographic sort (GEOSORT) in the Canadian Census, and what impact it may have on hot deck donor imputation results.

Imputation Classes by Size Measure for the Annual Survey of Manufacture of Statistics Canada

❖ Yi Li, Statistics Canada, 150 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *yi.li@statcan.ca*

Key Words: Ratio Type Imputation Methods, Imputation Classes, Size Measure

Ratio type imputation methods are extensively used for the imputation of the Annual Survey of Manufacture (ASM) of Statistics Canada. For example, one method consists in deriving the ratio between the variable requiring imputation and an auxiliary variable using a group of eligible records within an imputation class. This ratio is then applied to the record requiring imputation to generate an imputed value. In order to improve the quality of imputed data, in addition to industry classification and geographical location, a size measure variable was introduced in constructing the imputation classes so that the units of different sizes can be properly represented. In this paper, we will discuss the criteria used to define the size measure variable and the strategy used to determine the size groups for the Edit and Imputation of ASM.

289 Nonparametric Methods for Reliability and Health Sciences Applications ✪

Section on Nonparametric Statistics, Section on Health Policy Statistics, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 10:30 am–12:20 pm

A Nonparametric Test for Homogeneity of Survival Means

❖ Mohammad Rahbar, Michigan State University, B644 West Fee Hall, Epidemiology Department, East Lansing, MI 48824, *rahbar@msu.edu*; Sangchoon Jeon, Michigan State University; Joseph C. Gardiner, Michigan State University

Key Words: Nonparametric Methods, Survival Analysis, Cox Proportional Hazard Model, Right Censored Data, Kaplan-Meier Method, Simulations

Statistical procedures for assessment of treatments often involve complexities due to incomplete data. Here we propose a nonparametric method for testing the homogeneity of survival means based on right censored data. We develop a test based on Kaplan-Meier means. We derive asymptotic properties of this test statistic. Through simulations we compute and compare the power of this new procedure with that of the log-rank, Wilcoxon, and the Cox model. Our results indicate that performance of these estimation procedures depends on the level of censoring and whether the underlying assumptions are met. When the assumptions of log-rank and the Cox model are met, these procedures are more powerful than the other two tests. However, when the objective is testing homogeneity of survival mean, rather than survival curves, our new test statistic and Wilcoxon seems to have some advantages.

The Simultaneous Use of Weighted Logrank and Weighted Kaplan-Meier Statistics with Clustered Right-Censored Data

Yunchan Chi, National Cheng Kung University; ❖ Pei-Fang Su, National Cheng Kung University, Department of Statistics, 701 Taiwan, *spf@stat.ncku.edu.tw*

Key Words: clustered right-censored data, logrank test, Kaplan-Meier test

In observational studies, clustered right-censored data arise often. For example, to determine whether treatment extends the lifetime of ventilating tubes, the children with otitis media are randomly assigned into two groups. The lifetimes of tubes in both ears for a child (cluster) are correlated. For testing the equality of two survival distributions with such data, the weighted logrank test (WLR) and the weighted Kaplan-Meier test (WKM)

have been extended, respectively. The weights of both tests can be selected to be more sensitive to detect a certain alternative. However, one seldom knows what the exact alternative is. Therefore, this paper develop a class of versatile test based on the simultaneous use of WLR and WKM tests for clustered survival data. Furthermore, simulation studies show that the versatile tests maintain better power across a broad range of alternatives.

Exact, Distribution-Free Confidence Intervals for Late Effects in Censored Matched Pairs

❖Shoshana Daniel, University of Pennsylvania, 501 Waldron Terrrace, Merion, PA 19066, *skrieger@mail.med.upenn.edu*; Paul Rosenbaum, University of Pennsylvania

Key Words: late effects, matched pairs, permutation test, sign test, signed rank test

In a study of provider specialty in the treatment of ovarian cancer, a late divergence in the Kaplan-Meier survival curves hinted at superior survival among patients of gynecological oncologists when compared to patients of medical oncologists; we ask whether this late divergence should be taken seriously. Specifically, we develop exact, permutation tests, and exact confidence intervals formed by inverting the tests, for late effects in matched pairs subject to random but heterogeneous censoring without requiring knowledge of censoring times for patients who die. Exact distributions are consequences of three results about signs, signed ranks, and their conditional independence properties. A simulation shows that the late effects signed rank test has substantially more power to detect late effects than do conventional tests.

Use of the Spike-Motif Distance Metric

❖Katherine Tranbarger, Amherst College, Amherst College Dept of Math and CS, AC box 2291, Amherst, MA 01002, *ktranbarger@amherst.edu*

Key Words: point pattern, distance metric, prototype pattern, Basketball

One of three distance metrics proposed by Victor and Purpura for examining neuronal impulse trains, is the spike-motif distance. Past work has shown how a related distance (spike-time) can be used to determine a prototype sequence such that the distance between the prototype and all sequences in the data set is minimized. This work illustrates possible advantages of the spike-motif distance and how prototypes can be computed under this alternate metric. As an illustration, the method is shown applied to the timing of foul calls in professional Basketball.

Semiparametric Estimation of Hazard Function with Failure Indicators MAR

❖Chunling Liu, University of Hong Kong, Dept of Statistics and Actuarial Science, Hong Kong, HK Hong Kong, *cliu1221@hku.hk*; Qihua Wang, University of Hong Kong

Key Words: Hazard function, regression surrogate estimate, inverse probabilty weighted estimate, Missing at random, semi-parametric estimation

This paper presented three estimators for the hazard rate function under the semi-parametric right random censoring model with some failure indicators missing at random. The strong consistency and asymptotic normality properties were studied. The numerical simulation indicates the proposed estimators are superior to the nonparametric estimators.

Nonparametric Estimation of Mean Residual Life Function

❖Shufang Liu, North Carolina State University, 3520 Cum laude Ct, Apartment 301, Raleigh, NC 27606, *sliu@ncsu.edu*; Sujit Ghosh, North Carolina State University

Key Words: Mean residual life function, Right-censored data, Scale mixtures, Smooth nonparametric estimate

The mean residual life function (mrlf) of a subject is defined as the average residual lifetime of the subject given that the subject has survived up to a given time point. It is well known that under mild regularity conditions an mrlf completely determines the probability distribution of the subjects' lifetime. In practice, the advantage of the mrlf over the more popularly used hazard function lies in its interpretation in many applications where the primary goal is often to characterize the remaining life expectancy of a subject instead of the instantaneous failure rate. A smooth nonparametric estimator of the mrlf is proposed using scale mixtures of the empirical estimate of the mrlf based on right-censored data. Asymptotic properties are established. Empirical performances of the proposed estimator are studied based on simulated data sets and a real dataset.

Two-Fold Heteroscedastic Nested Models When the Number of Subclasses Is Large

❖Shu-min Liao, The Pennsylvania State University, 326 Thomas Building, Department of Statistics, University Park, PA 16802, *sxl340@psu.edu*; Michael G. Akritas, The Pennsylvania State University

Key Words: Nonparametric, Nested models, Heteroscedastic, Subclass effects, Large number of subclasses

The purpose of this talk is to present our latest investigation about two-fold nonparametric nested models. The testing of the subclass effects when the number of subclasses is large is the core of this talk. We start from the homoscedastic balanced cases, and then move on to the heteroscedastic unbalanced cases, in which two different approaches are proposed and compared. Not only the fixed-effects models but also the random-effects and mixed-effects models are studied.

290 Helpful Teaching Techniques

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 10:30 am–12:20 pm

An Experience in Teaching Inferential Statistics with Focus on the Analyses of Simple Survey Data

❖Bodapati V.R. Gandhi, University of Puerto Rico, Calle Azorin 710, Mansiones de Espana, Mayaguez, PR 00682, *bvrgandhi@yahoo.com*

Key Words: survey data, inferential statistics, teaching, cooperative learning, business statistics, statistical education

The paper focuses on teaching inferential statistics to business students via data analysis of simple survey information. In general, students are taught various techniques to solve textbook problems in business statistics courses. However, students are often unclear about how to establish the proper propositions to be tested and the formulation of the relevant null hypotheses in real life situations in surveys. The paper narrates the teaching experience when students were required to administer and analyze their own questionnaire data in an environment of cooperative learning. The author concludes with recommendations to increase the effectiveness of statistics teaching to better serve the business curriculum.

Teaching Data Mining to Undergraduate Business Students

❖Margaretha Hsu, Shippensburg University, 1871 Old Main Dr, Shippensburg, PA 17257, *mmhsu@ship.edu*

Key Words: data mining, statistics, data analysis, business

Traditional statistic courses are taught in almost all business schools. They usually cover the more classic statistical methods. The data used in these courses also tend to be small, well balanced, and organized. With the advances of computer capability, it becomes much easier for students to deal with large amounts of data, and experience the fun and the power of data analysis. The tools students learn in data mining can be applied to a wide range of areas which include banking, health care, market research, manufacturing, and retailing. In order to fully appreciate the results of data mining, students may need a certain level of statistical education. This article discusses the benefits and challenges of teaching data mining to undergraduate business students who have a minimum knowledge of statistics.

Students Don't Do 'Optional'

❖Kimberly Denley, University of Mississippi, 217 Hume Hall, University, MS 38677, *kdenley@olemiss.edu*

In an attempt to implement the maxim "doing it is better than watching it," the University of Mississippi introduced computer aided learning into the classroom experience. But to gain maximum results, they had to realize that "students don't do optional." The presentation will present results from a hybrid course redesign in elementary statistics.

Teaching Hypothesis Testing Using a Magic Trick

❖Gregory L. Snow, Intermountain Healthcare, 8th Avenue and C street, Salt Lake City, UT 84143, *greg.snow@intermountainmail.org*

Key Words: hypothesis testing, teaching, examples, magic

One of the hardest concepts for statistics students to understand is hypothesis testing. Most students find it counter-intuitive and don't find the common examples to be intuitive. I have found that introducing or reviewing the concept using a simple magic trick can help students to understand hypothesis testing in a more intuitive way. I will demonstrate the trick along with how I use it in the classroom.

Pig Data and Bayesian Inference on Multinomial Probabilities

❖John Kern, Duquesne University, 600 Forbes Avenue, 440 College Hall, Pittsburgh, PA 15282, *kern@mathcs.duq.edu*

Key Words: multinomial, Dirichlet, posterior predictive distribution

Bayesian inference on multinomial probabilities is conducted based on data collected from the game Pass the PigsÆ. Prior information on these probabilities is readily available from the instruction manual, and is easily incorporated in a Dirichlet prior. Posterior analysis of the scoring probabilities quantifies the discrepancy between empirical and prior estimates, and yields posterior predictive simulations used to compare competing extreme strategies.

Using Humming To Teach Experimental Design in Introduction Statistics

❖Cora Neal, Sonoma State University, 1801 East Cotati Ave, Darwin 114, Rohnert Park, CA 94928, *cora.neal@sonoma.edu*

Key Words: teaching, introductory statistics, experimental design, hands-on

This interactive presentation will demonstrate a technique that has been successfully used to teach experimental design concepts to introductory statistics students at the college level. By starting the class with this simple activity a formerly boring lecture filled with terms and definitions has been transformed to a lively discussion where students generate the ideas and the instructor attaches the appropriate vocabulary. Audience participation is required, so come ready to hum.

Using Matlab in the First Probability and Statistics Sequence

❖John Turner, U.S. Naval Academy, 225 Autumn Ln, Centreville, MD 21617, *jct@usna.edu*

Key Words: order statistics, programming, confidence intervals

For a variety of reasons, we chose Matlab for use in our first (two semester) sequence in probability and statistics. This has had several impacts on this course. Programming occurs naturally and in such a way that it encourages the students to customize their routines. However, the biggest impact has been on the topics that can be covered and how they are covered. This has impacted both confidence intervals and the use of order statistics in the course.

291 Environmental Health Effects

Section on Statistics and the Environment, ENAR, Section on Health Policy Statistics, Section on Statistical Graphics

Tuesday, July 31, 10:30 am–12:20 pm

Targeting Areas at Risk for Childhood Lead Poisoning

❖Margaret Conomos, U.S. Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, 2842T, Washington, DC 20460, *Conomos.Margaret@epamail.epa.gov*; Warren Strauss, Battelle; Tim Pivetz, Battelle; Jyothi Nagaraja, Battelle; Elizabeth Sloan, Battelle; Rona Boehm, Battelle; Michael Schlatt, Battelle; Darlene Wells, Battelle; Michele Morara, Battelle; Bruce Buxton, Battelle; Barry D. Nussbaum, U.S. Environmental Protection Agency; Heather Case, U.S. Environmental Protection Agency

Key Words: childhood lead poisoning, targeting geographic areas, visualization tool, risk assessment, statistical modeling, GLMM

This pilot study seeks to develop statistical models to predict risk of childhood lead poisoning within specified geographic areas, based on a combination of demographic, environmental, and programmatic information sources. Exposure factors associated with childhood lead poisoning were investigated within census tracts for a community-focused set of Models in Massachusetts, as well as within counties across the US in a series of National Models. Aggregated summary measures, geometric means, and proportions of children screened for lead poisoning within defined geographic areas are being used as response variables in the statistical models. These summary measures are reported at 3-month intervals over several yearly time periods within each geographic area, allowing EPA to assess how the risk of childhood lead poisoning changes over time using a Generalized Linear Mixed Modeling Approach.

Spatio-Temporal Dynamics of Eastern Equine Encephalitis in Southern New Hampshire

❖Ernst Linder, University of New Hampshire, Dept of Mathematics and Statistics, Kingsbury Hall, Durham, NH 03824, *elinder@unh.edu*; Carlee Moser, University of New Hampshire; Francois Laflamme, Statistics Canada; Jason Stull, U.S. Department of Health & Human Services

Key Words: spatial surveillance, cluster detection, space time modeling, sparse data, disease mapping

Eastern Equine Encephalitis (EEE) is a mosquito transmitted virus that is deadly to horses and can be deadly to humans. It has been prevalent during the summer in the Eastern US. Recent equine and human cases of EEE in the Southern portion of New Hampshire have prompted the pooled testing for EEE of numerous mosquito samples. We examine the spatio temporal dynamics of the occurrence of EEE based on such mosquito pool data for 37 towns using weekly aggregated counts during the last three summers. We pose a spatio temporal model for p, the probability that a randomly selected mosquito is a carrier. We assume a functional prior for the mean of p over time that reflects the onset, rise, and decline of the virus over a typical season. Our model allows us to compare occurrence rates between towns but also to identify clusters or hot spots of elevated risk in space and time.

Exploring the Relation Between Frailty and Negative Health Outcomes Associated with Air Pollution in an Older Population

❖ Sandrah Eckel, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, 615 N Wolfe St Room E3527, Baltimore, MD 21205, *seckel@jhsph.edu*; Thomas A. Louis, Johns Hopkins University; Karen Bandeen-Roche, Johns Hopkins Bloomberg School of Public Health; Linda Fried, Johns Hopkins Medical Institutions; Helene Margolis, California Department of Health Services

Key Words: effect modification, frailty models, longitudinal models

In relating air pollution exposure to health effects in older adults, age is a noisy surrogate of health status. In gerontology, an individual is frail if reduced to a vulnerable state and at risk for increased disability or disease. Our goal is to implement models using the frailty concept in place of age. We develop, evaluate and apply statistical methodology for longitudinally analyzing the health effects of an exposure (air pollution) while taking into account the heterogeneity of individual level health status, both across time and between individuals. We use data from the Cardiovascular Health Study (CHS) and the Environmental Factors Ancillary Study to CHS. Statistical challenges include representing frailty in our model, developing appropriate predictor summaries and choosing appropriate longitudinal and survival model frameworks.

Analysis of the Effect of Pollutants on Respiratory Infection Using Case-Crossover Design and Generalized Additive Models

❖ Long H. Ngo, BIDMC & Harvard Medical School, 1309 Beacon Street, CO 203, Brookline, MA 02446, *long2ucb@gmail.com*; Dzung V. Do, University of Medicine and Pharmacy; Thuan Q. Thach, University of Hong Kong; Sumi Mehta, Health Effects Institute; Aaron J. Cohen, Health Effects Institute

Key Words: pollutants, ALRI, case-crossover, Poisson regression, generalized additive mixed model

In a study to evaluate the effect of ambient pollutants (e.g., PM10) on acute lower respiratory infection (ALRI) in the children of Ho Chi Minh City in Vietnam, we used case-crossover and Poisson regression to analyze the data from 2003 to 2005. We evaluated both the symmetric bi-directional design, and time-stratified design for the case-crossover analysis. Generalized additive mixed models were used for the Poisson models. Systematic missing data of the pollutants and measurement error due to misclassification of diagnosis and exposure estimated from the central monitors, and multicollinearity among the pollutants posed problems for inference. We hypothesized that poverty status, using treatment fee, acted as a modifier on the association between pollutants and ALRI. We explored the issues of non-random missingness, small sample size, and misclassification in both methods of analysis.

Demographic and Behavioral Modifiers of Arsenic Exposure Pathways: A Bayesian Hierarchical Analysis of NHEXAS Data

❖ Rajib Paul, The Ohio State University, 101 Curl Drive, Room No. 647, Columbus, OH 43210, *rajib@stat.ohio-state.edu*; Catherine A. Calder, The Ohio State University; Peter F. Craigmile, The Ohio State University; Thomas Santner, The Ohio State University

Key Words: Biomarkers, Censored data, EPA Region 5, MCMC, Missing data, Subpopulation analysis

We describe a Bayesian hierarchical model that explains subpopulation-specific pathways of exposure to arsenic from sources in the environment to excretion from the human body. Our model is fitted to data collected as part of the National Human Exposure Assessment Survey (NHEXAS). Using demographic information (e.g., gender, age), as well as surrogates for environmental exposure (e.g., usage of gas equipment, average number of minutes spent in an enclosed workshop), we identify subpopulation differences in exposure routes. Missing and censored data are handled systematically in the Bayesian framework, and inferences on model parameters are obtained using a Markov chain Monte Carlo (MCMC) algorithm.

A Correlation Analysis of Bivariate Normal Data Subject to Partial Left-Censoring

❖ Wendy Leith, Portland State University, Department of Mathematics and Statistics, PO Box 751, Portland, OR 97207, *leith@pdx.edu*; Jong Kim, Portland State University; Tom Fielden, Portland State University

Key Words: EM algorithm, left-censored data, bivariate normal, bivariate lognormal, correlation

Researchers working with bivariate normal data subject to partial left-censoring have not had access to software which could treat the data appropriately. In an effort to treat the data appropriately and to fully utilize the information the censored observations provide we use the EM Algorithm to obtain the MLEs of the parameters. After arriving at a final estimate we use Monte Carlo Imputation to estimate the standard errors. Finally, we extend our work to cover partially left-censored bivariate lognormal data, which is commonly studied in environmental sciences. Simulations show that the EM algorithm is an excellent choice for estimating the parameters of the bivariate normal distribution when both variables are subject to partial left censoring, and that the method is robust. The estimates perform very well regardless of sample size, strength of correlation or amount of censoring.

292 Statistical Issues in Case-control and Combined Study Designs ●

Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 10:30 am–12:20 pm

Sample-Size Calculations in Case-Control Unmatched Studies When Controls Contain Cases

❖ Xiaolei Liu, University of Cincinnati, Environmental Health Dept, Cincinnati, OH 45220, *liuxo@email.uc.edu*; Marepalli B. Rao, University of Cincinnati

Key Words: Case-control, Misclassification, Sample size, Sleep Apnea, Contingency table

In this paper, the focus is on 2◊2 case-control unmatched studies, in which the cases (those with the disease) are genuine but the controls may contain patients with disease due to misclassification in disease screening. We systemically examine the effect of such contamination on odds ratio and data analysis. We will provide sample size formulas using two different procedures by taking into account the contamination ratio, one procedure is by testing equivalence of two sample proportions and the other using log odds ratio test. The misclassification problem arose when planning a study to identify factors causing sleep apnea.

Weighting for Case-Control Studies with Partial Matched Controls

❖Ronaldo Iachan, Macro International, 20705 Beltsville Dr, Calverton, MD 20705, *ronaldo.iachan@orcmacro.com*; Aliza Fink, Macro International

This analysis was motivated by a matched case control study where controls were not recruited for a large proportion of the cases. An analytic solution was explored that capitalized on the known distribution of the matching factors in the cancer registry population used to select both cases and controls. The method applied post-stratification weights to the participants to make the weighted sample reflect the total eligible population. This analysis will determine whether this method can reduce the selection bias created by matching. Our simulation study will use data on case/control status and matching factors taken from the state cancer registry. The analysis, comparing odds ratios, will be repeated for varying a) response rates and ns (controls per case), b) the association between the exposure and the outcome, c) the association between the exposure and the matching factors.

A Latent Group Approach for Combining Matched and Unmatched Case-Control Studies

❖Mulugeta Gebregziabher, MUSC, 135 Cannon St suite 303, Charleston, SC 29425, *gebregz@musc.edu*; Paulo Guimaraes, MUSC; Wendy Cozen, University of California, Los Angeles; David Conti, University of Southern California

Key Words: case-control, conditional likelihood, latent group, polytomous conditional likelihood

In case-control studies, it is common practice to compare two or more different sets of controls with the same case group to validate or confirm a positive or negative finding. This usually involves fitting separate models for each case-control comparison, testing the homogeneity of the parameters and if appropriate obtaining a pooled estimate. But, fitting separate models tends to lead to a homogeneity test with inflated type-II-error rate and to a pooled estimate with larger standard error. The problem is compounded when one control group is matched and the other is unmatched. The alternative to fitting separate models is to use a multinomial model. However, available methods for combining matched and unmatched case-control data do not handle multinomial response. We propose a unified latent group approach that can be used for both binary and multinomial response case-control data.

Joint Modeling for Survival and Longitudinal Data Under Nested Case-Control Sampling

❖Chi-hong Tseng, New York University, 650 1st Ave 540, Medical Center, New York, NY 10016, *tsengc01@med.nyu.edu*; Mengling Liu, New York University

Key Words: nested case-control, of Cox proportional hazards model, maximum likelihood estimator, measurement error

Nested case-control study design is popular in epidemiological practice due to its cost-effectiveness. Typically, for all the cases and the selected (matched) controls, complete information is collected, while the information on other members in the study cohort is limited. The ordinary conditional logistic regression is widely used, however, its efficiency can be of concern. We present a maximum likelihood estimator under the nested case-control sampling in the context of Cox proportional hazards model. In particular, we are interested in estimating the effect of a time-dependent exposure that have a variable number of repeated measurements in cases and their matched controls. The method is applied to data from a nested case-control study of hormones and breast cancer.

Statistical Methods for Detecting Microdeletions in Genetic Markers

❖Chih-Chieh Wu, The University of Texas M.D. Anderson Cancer Center, Department of Epidemiology Unit 1340, 1155 Pressler Street, Houston, TX 77030, *ccwu@mdanderson.org*; Sanjay Shete, The University of Texas M.D. Anderson Cancer Center; Christopher I. Amos, The University of Texas M.D. Anderson Cancer Center

Key Words: heterozygosity, homzygosity, microdeletion, Pearson's Chi-Squared test, run statistic, statistical test

There is increasing evidence that many human genetic disorders are associated with deletions of DNA sequences in chromosomes. Identified microdeletions range in size from < 1kb to 4Mb. Because the current molecular techniques are not efficient for direct identification of genetic deletions, we have developed marker-based statistical methods for detecting regions of microdeletion using case-control designs. We propose to test for an excess of homozygosity of contiguous genetic markers in patients presenting with a genetic disorder. Our approaches evaluate the frequency of contiguous homozygosity segments in a subject and determine the heterozygosity level at each region among subjects. The sensitivity and power of our proposed methods depend on marker density, allele frequency, and size of microdeletion. We will perform analyses for simulation studies and lung cancer data.

Estimating the Prevalence of Disease Using Relatives of Case and Control Profanes

❖Kristin Javaras, Harvard School of Public Health, Department of Biostatistics, 655 Huntington Ave Bldg II Fourth Fl., Boston, MA 02115, *kjavaras@hsph.harvard.edu*; Nan Laird, Harvard University; Brian D. Ripley, Oxford University; James I. Hudson, Harvard Medical School

Key Words: Disease prevalence, Case-control family study, Proportion

Case-control family studies are conducted to investigate familial aggregation of a disease. However, the resulting data (disease status and covariates for members of families selected via case and control probands) can also be used to estimate the prevalence of the disease. We propose estimators for overall and covariate-stratum-specific prevalence that are approximately unbiased for their population counterparts, given certain commonly made assumptions. We also propose corresponding confidence intervals that have good coverage properties, even for small prevalence's. The estimators and intervals address the over-representation of diseased individuals by excluding profanes and taking into account whether each relative was selected via a case or control probed. The estimators and intervals perform well in simulation experiments designed to mimic a real family study of major depression.

Alleviating Ecological Bias in Generalized Linear Models with Optimal Subsample Design

❖Adam N. Glynn, Harvard University, 26 Cambria St, Unit 2, Somerville, MA 02143, *aglynn@iq.harvard.edu*; Jon Wakefield, University of Washington; Mark S. Handcock, University of Washington; Thomas Richardson, University of Washington

Key Words: Ecological bias, Combining information, Within-area confounding, Sampling

In this paper, we illustrate that combining ecological data with subsample data in situations in which a generalized linear model (GLM) is appropriate provides two main benefits. First, by including the individual level subsample data, the biases associated with ecological inference in GLMs can be eliminated. Second, we can use readily available ecological data to design optimal subsampling schemes, so as to maximize information about parameters. We present applications of this methodology in sociology and epidemiology, showing that small, optimally chosen subsamples can be combined with ecological data to generate precise estimates relative to a simple random subsample.

293 Survey Response Propensity and Responsive Design ●

Section on Survey Research Methods

Tuesday, July 31, 10:30 am–12:20 pm

Evaluating Response Quality of Nonrespondents Using the Imputation Technique

❖ Cong Ye, University of Maryland, JPSM, 1218 Lefrak Hall, College Park, MD 20742, *eyecong@msn.com*; Ting Yan, University of Michigan; Mandi Yu, University of Michigan

Key Words: Measurement Error, Missing Data, Imputation, Response Propensity

Declining response rates in household surveys speak to the increasing difficulty in recruiting sampled persons to respond to a survey request. Survey organizations are spending increasingly more effort on contacting and persuading respondents with a low response propensity in an attempt to increase response rates and reduce nonresponse bias. A major concern, however, is that, if respondents with a low response propensity turned out to be bad reporters, the costly extra recruitment effort could reduce nonresponse error at the expense of increasing measurement error. However, this concern is difficult to address since the measurement error property of nonrespondents is unknown. This paper treats this issue as a missing data problem and imputes for nonrespondents hard to recruit. The preliminary results showed that respondents who are hard to recruit didn't provide worse data.

Using Callback Models To Adjust for Nonignorable Nonresponse in Face-to-Face Surveys

❖ Paul Biemer, RTI International/The University of North Carolina at Chapel Hill, PO Box 12194, Research Triangle Park, NC 27709, *ppb@rti.org*; Kevin Wang, RTI International

Key Words: paradata, latent class analysis, drug survey, weighting class adjustment

Traditional methods of adjusting for survey nonresponse bias assume nonresponse is "ignorable." Biemer and Link (2006) propose an alternative approach that relaxes this assumption by using so-called level of effort (LOE) variables derived from call attempts to model the response propensity. Using a latent indicator variable, the model distinguishes between sample members who will eventually respond to a survey with sufficient call attempts and those that will never respond regardless of the number of call attempts (i.e., the "hard core" nonrespondents). They applied their callback models to a large RDD survey. The present paper applies a similar model to a face to face survey, viz., the National Survey of Drug Use and Health (NSDUH). The nonresponse bias reduction of this approach is assessed and compared with that of the traditional approach as implemented in the NSDUH.

Assessing Bias in Estimates in a Two-Stage Design from an Early Close Out of the First Stage Data Collection: An Empirical Investigation Using NSRCG Sample Data

❖ Donsig Jang, Mathematica Policy Research, Inc., 600 Maryland Ave, SW Suite 550, Washington, DC 20024, *don_sig@yahoo.com*; Xiaojing Lin, Mathematica Policy Research, Inc.

Key Words: Nonresponse analysis, NSRCG, School sample, Bias

Reluctant respondents and low response rates have resulted in increased data collection costs to maintain the same level of response from one year to the next. Consequently, survey managers must assess the efficient allocation of a fixed budget to achieve the survey objectives when the survey is conducted. The list collection of college graduates is a major component of the NSRCG design and one that has considerable costs associated with it. In particular, data collection resources are concentrated on a small set of late responding schools. These resources could be used elsewhere if the schools responded earlier. This paper focuses on the effect of school nonresponse if the list collection period was not extended and a higher school level nonresponse rate was accepted. With this objective, we assess the bias of survey estimates due to school-level nonresponse at varying response rates.

Optimal Survey Design When Nonrespondents Are Subsampled for Follow-Up

❖ Alistair J. O'Malley, Harvard Medical School, Department of Health Care Policy, 180 Longwood Avenue, Boston, MA 02115-5899, *omalley@hcp.med.harvard.edu*; Alan M. Zaslavsky, Harvard University

Key Words: Neyman allocation, Nonrespondents, Small Area Estimation, Subsample, Survey, Telephone Followup

Healthcare surveys often first mail questionnaires to sampled members of health plans and then follow up mail nonrespondents by phone. The high unit costs of telephone interviews make it cost-effective to subsample the follow-up. We derive optimal subsampling rates for the phone subsample for comparison of health plans. Computations under design-based inference depart from the traditional formulae for Neyman allocation because the phone sample size at each plan is constrained by the number of mail nonrespondents and multiple plans are subject to a single cost constraint. Because plan means for mail respondents are highly correlated with those for phone respondents, more precise estimates (at fixed overall cost) for potential phone respondents are obtained by combining the direct estimates from phone follow-up with predictions from the mail survey using small-area estimation (SAE) models.

Research and Responsive Design Options for Survey Data Collection at Statistics Canada

❖ Chris Mohl, Statistics Canada, 17M RH Coats Bldg, 150 Tunneys Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *chris.mohl@statcan.ca*; Francois Laflamme, Statistics Canada

Key Words: Paradata, Responsive Design, Collection Procedures

In an effort to continue to collect timely, high quality, but cost-effective survey data, Statistics Canada is reviewing its current data collection practices. The use of paradata is seen as a key component of this analysis. This information could be used to implement a responsive design which would allow the collection strategy to evolve over time. The paradata can be used either on its own or in conjunction with historical information to provide clues on when initial or subsequent contact attempts should be made, assign differing collection priorities to units or sub-sample non-respondents late in the collection cycle. The options being studied as part of this research project are discussed and some initial findings are presented. Possible responsive design steps, their benefits and the information needed to implement them are also discussed.

A Dynamic Approach to Response Propensity

❖Kristen Olson, University of Michigan, Institute for Social Research, 426 Thompson St Room 4050, Ann Arbor, MI 48104-2321, *olsok@isr.umich.edu*

Key Words: nonresponse, response propensity, nonresponse bias, data collection

Traditional views of survey participation view a person's likelihood of participating in a survey as deterministic or stochastic. Both views assert that a sample unit has one response propensity, given an evolved survey protocol. In a field survey, sample units vary in their receipt of individual recruitment protocol components and in the timing of these components. These components are specifically applied to alter response propensities. This paper argues that sample units have more than one response propensity, changing with the introduction of new protocol components. We call this a dynamic view of response propensity. Dynamic response propensities are estimated using two surveys with records for respondents and nonrespondents. The records are used to illustrate how a dynamic response propensity approach changes the understanding of nonresponse bias and of recruitment procedures.

Findings from Methodology Innovations on Population Health Survey

❖Urban Landreman, Hennepin County HSPHD, 525 Portland Avenue S L963, Minneapolis, MN 55415, *urban.landreman@co.hennepin.mn.us*

Key Words: DSF, health, mixed-mode, CATI, GIS, bias

The 2006 SHAPE health survey for adults & children for Hennepin Cnty., MN used several innovations in methodology for better coverage and response: Using USPS Delivery Sequence File along with an all-listed phone number sample; Mail along with telephone CATI mode; and using 'nearest neighbor' replacement when primary household in sample was unavailable. The innovations allowed for analysis of quality of listed phone number sample frame compared to county population; mode effects and sampling bias on health questions; and demonstration of feasibility and benefits of 'nearest neighbor' technique with DSF.

294 Contributed Poster Presentations

Business and Economics Statistics Section, Section on Bayesian Statistical Science, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Quality and Productivity, Section on Risk Analysis, Section on Statistical Consulting, Section on Statistical Education, Section on Teaching Statistics in the Health Sciences, IMS, Section on Statisticians in Defense and National Security

Tuesday, July 31, 10:30 am–12:20 pm

Assessment Meets Institutional Research

❖Julia Norton, California State University, East Bay, 28022 El Portal Drive, Hayward, CA 94542, *julia.norton@csueastbay.edu*

Key Words: assessment, Deparments of Statistics, Institutional Research

Paper analyzes FROSH experience at a medium sized western state university and the relationship with faculty established General Education Learning Outcomes. Suggestions are considered for establishing baselines and tying expectations with coursework. Do positive outcomes result in higher retention rates f? What effect does advising have on student retention? Suggestions for carrying over FROSH good will into the second college year experience are considered. Learning communities are studied as a means for linking learning with expectations.

Bayesian Analysis of Correlated Proportions (Matched-Pair Data) in 2 x 2 Tables with Incompletely Classified Data: Bayesian Extension of McNemar's Test to Missing Data

❖Yan Lin, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd Unit 447, Houston, TX 77030, *hyanlin@mdanderson.org*; Stuart R. Lipsitz, Brigham and Women's Hospital; Debajyoti Sinha, Medical University of South; Scott Regenbogen, Harvard School of Public Health; Caprice Greenberg, Brigham and Women's Hospital

Key Words: Dirichlet prior, Gibbs sampler, Ignorable missing data, 2 x 2 table

Altham (1971) proposed a Bayesian analysis of a 2 x 2 contingency table formed from matched pairs. Using the same Bayesian perspective, we develop an extension of Altham's Bayesian methods to a 2 x 2 table from matched pairs with missing data and under an ignorable missing data mechanism. In particular, we develop a Bayesian p-value similar to the p-value from McNemar's test, which, unlike McNemar's test, incorporates the missing data into the analysis. We use a simulation study to explore the power and Type I error rate of our proposed test statistic.

A Bayesian Model for Inversion of Geophysical Seismic and Electromagnetic Data

❖Jinsong Chen, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, 90R1116, Berkeley, CA 94720, *jchen@lbl.gov*; Susan Hubbard, Lawrence Berkeley National Laboratory; Michael Hoversten, Chevron Energy Technology Company; Alan Hubbard, University of California, Berkeley

Key Words: Bayesian model, MCMC, Joint inversion, Geophysics, Parameter estimation, Slice sampling

Reliable estimates of gas saturation need to combine the commonly used seismic imaging techniques with the recently developed controlled-source electromagnetic (CSEM) methods. However, conventional methods for inversion of seismic and CSEM data are limited because such methods depends on the choice of the initial model parameters and cannot provide adequate information about uncertainty in the estimation. In this study, we develop a Bayesian model to jointly invert seismic and CSEM data and use a mixed sampling method that includes single and multivariate Metropolis-Hastings and slice sampling methods to explore the joint probability distribution function. Synthetic and field case studies show that the developed Bayesian model is effective for integration of seismic and CSEM data resulting in better estimates of gas estimation.

Probabilistic Plan Tracking and Detection for Intelligence Analysis

❖Sinjini Mitra, The University of Southern California, Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292, *mitra@isi.edu*; Paul Cohen, The University of Southern California; Aram Galstyan, The University of Southern California

Key Words: Hidden Markov Model, plan recognition, tracking, belief state

Plan recognition is the problem of inferring an agent's unobservable state of plans or intentions based on observations of its interaction with the environment. In this paper, we present a rigorous theoretical framework for detecting and tracking malicious plans based on Abstract Hidden Markov Models that uses a Dynamic Bayesian Network representation of the plan hierarchy. The problem is to determine the top-level policy along with those at the lower levels given the current sequence of observations by updating the belief state using the posterior distribution at each time point. Unlike most existing methods, our model does not assume the known identity of the agent and is capable of tracking a very big network of agents carrying out different types of transactions and accurately detecting groups that intend to cause harm. We test our method on a virtual society of agents, called Hats.

Approximations to Continuous Dynamic Processes in Hierarchical Models

❖ Amanda R. Cangelosi, Utah State University, 3900 Old Main Hill, Logan, UT 84321, *acang@cc.usu.edu*; Mevin Hooten, Utah State University

Key Words: Hierarchical models, Bayesian statistics, Dynamical processes, Population dynamics

Modeling nonlinear processes, such as population dynamics, has been given much attention in mathematics. For example, species competition has been extensively modeled by differential equations. It is of both scientific and mathematical interest to utilize such models in a statistical framework. This incorporation allows one to quantify uncertainty associated with random effects, such as demographic and environmental stochasticity, and thus assess the accuracy of deterministic model components. In this pursuit, parameters of the process models are characterized in a Bayesian hierarchical context, using discretizations with varying levels of precision to assess the necessity of higher-order approximations.

Bayesian Modeling of Ecosystems in the Permian Mass Extinction

❖ Michael Karcher, Swarthmore College, 2436 North Line Street, Colmar, PA 18915, *mkarche1@swarthmore.edu*; Steve Wang, Swarthmore College; Peter Roopnarine, California Academy of Sciences; Kenneth D. Angielczyk, California Academy of Sciences

Key Words: paleontology, MCMC, simulation-based inference, food webs, Metropolis-Hastings

The Permian mass extinction was the most severe extinction event in the history of life, killing an estimated 90% of all species. However, its causes remain mysterious. Using fossil data from the Karoo Basin in South Africa, we construct a food web model of Permian ecosystems before the extinction event. We group fossil species into 22 guilds based on their ecological roles and use a paleobiological computer model to simulate perturbations to these food webs. We then use Markov Chain Monte Carlo techniques, specifically the Metropolis-Hastings Algorithm, to construct a Bayesian posterior distribution for the various extinction-event scenarios, thereby inferring the initial level of extinction in each guild.

A Bayesian Nonparametric Multivariate Model for Evaluation of Correlated Diagnostic Tests

❖ Adam Branscum, University of Kentucky, College of Public Health, 121 Washington Ave, Lexington, KY 40536, *abran3@email.uky.edu*; Timothy E. Hanson, The University of Minnesota; Ian Gardner, University of California, Davis

Key Words: Bayesian nonparametrics, Polya trees, Diagnostic tests

We develop a flexible nonparametric model for the statistical analysis of multivariate serologic data. In this setting, sampled individuals are tested for a condition of interest using several imperfect continuous diagnostic tests. To provide for a degree of robustness to the structure imposed by multivariate normal models, the data for diseased and non-diseased individuals are modeled with independent multivariate mixtures of Polya trees. Data-driven estimates are available for ROC curves and areas under them for all diagnostic tests under consideration. We consider a tractable two-stage empirical Bayesian estimation procedure that is applied to the evaluation of two ELISA tests for Johne's disease.

Pilot Study on Motivating High School AP Students To Learn Conditional Probability

❖ Neale Hirsh, NSH & Associates, 6619 South Dixie Hwy, 364, Coral Gables, FL 33143, *nhirsh2@mac.com*

Key Words: conditional, probability, Bayes, naturalistic, teaching

High school students seldom perform well unless they are directly motivated by the topic and the presenter. Conditional probability (CP) within AP Statistics is such a topic that's easy to overlook and assume it's understood, which usually is not true. In this naturalistic study of student performance, the dependent variables are the promptness and quality of homework submitted while CP is taught. We show that in-class games such as checkers and word games utilizing conditional probability influence both homework variables. Numerous uncontrolled variables exist, and are identified. Other experimental issues include the Hawthorne Effect and homogeneity and representativeness of the student population. Finally, with caution, we make educated guesses about what constitutes motivation for these students.

Bayesian Screening for Pharmacogenetic Effects in Clinical Trials

❖ Mengye Guo, University of Pennsylvania, 503 Blockley Hall
423 Guardian Drive, Department of Biostatistics & Epidemiology, Philadelphia, 19104, *mengyego@hotmail.com*; Daniel F. Heitjan, University of Pennsylvania

Key Words: Bayes factor, importance sampling, bupropion, pharmacogenomics, single-nucleotide polymorphism, Laplace approximation

Pharmacogenetics is concerned with the detection of genetic markers that modify treatment effects. Thus, statistical methods for pharmacogenetics aim to assess the significance of treatment-by-marker interactions. When the number of potential markers is large we encounter the problem of multiplicity. We propose a Bayesian hypothesis testing method (Berger, 1985) to screen a large pool of markers for statistically significant interactions; the method evaluates evidence using Bayes factors. We carry out the computations using both importance sampling and analytical approximation. The Bayesian method explicitly incorporates prior information. Moreover, it is less conservative than a frequentist test with a Bonferroni-type correction. We apply our method to a randomized trial of pharmacotherapy for smoking cessation, in which 84 SNPs were evaluated as potential pharmacogenetic markers.

Bayesian Statistical Power in Randomized Small Clinical Trials

❖ Yufeng Li, The University of Alabama at Birmingham, 1802 6th Avenue south, NP2540, Birmingham, AL 35216, *yufengli@uab.edu*; Shouluan Ding, The University of Alabama at Birmingham; Seng-jaw Soong, The University of Alabama at Birmingham

Key Words: Randomized clinical trial, Bayesian method, frequentist method;, Statistical power

Due to the small number of patients affected by some rare diseases, clinical trials studying these diseases are either underpowered because of the difficulty of recruiting adequate numbers of subjects, or the study recruiting period is prolonged, resulting in great expense. An appropriately designed trial is imperative. Currently, however, the methodologies for studying rare infectious diseases are not significantly different from the methodologies for studying other diseases. In this research, we study the Bayesian statistical approach in small randomized clinical trial with comparisons of the statistical power from Bayesian approach with frequentist approach. The results are compared using a simulated data.

A Stochastic Gravitational Field Simulated by a Monte Carlo Method: An Alternate Explanation for the Dynamics of Neutral Dust Grains in Saturn's Rings

❖Elizabeth Martinez-Gomez, Heriberto Frias 238-1, Col. Narvarte, C.P. 03020, Del. Benito Juarez, 04510 Mexico, *affabeca@gmail.com*

Key Words: Monte Carlo method, stochastic process, dust grains, Keplerian motion, gravitational field, power-law distribution

Saturn's ring system is the largest, brightest, and diverse in our solar system. For these reasons, it is also the most studied. The dusty particles in the rings are mainly composed of water-ice; their sizes follow a power-law distribution, and they orbit Saturn in a Keplerian motion. On the other hand, it is well-known that sputtering by charged particles and meteoroid bombardment of material from the rings create an extensive neutral cloud of water products in the inner magnetosphere. To explain the dynamics of the dust grains in Saturn's rings, a model is developed in which the gravitational field between neutral dust grains has a stochastic behavior simulated by a Monte Carlo Method. Our results show that the grains tend to gain energy over the time. Three energization regions are identified: a) High (Cassini Division, A-ring); b) Low (B and E rings), and c) Intermediate (C-ring).

Bivariate Generalized Poisson Regression Model

❖Felix Famoye, Central Michigan University, Department of Mathematics, Mt Pleasant, MI 48859-0001, *felix.famoye@cmich.edu*

Key Words: correlated count data, over-dispersion, goodness of fit, estimation

The bivariate generalized Poisson distribution has been found useful in describing correlated count data. In this paper, a bivariate generalized Poisson regression model based on the bivariate generalized Poisson distribution is defined and studied. The marginal means of the bivariate model are functions of the explanatory variables. The parameters of the regression model are estimated by using the maximum likelihood method. Some test statistics including goodness of fit are discussed. A numerical data that exhibits over-dispersion is used to illustrate the techniques.

The Application of Statistical Process Control to Chemical Solution Stability

❖Jennifer Herberich, Pratt & Whitney Rocketdyne, 1200 Rio Vista Blvd, Apt 205, Palm Beach Gardens, FL 33410, *jennifer.herberich@pw.utc.com*

Key Words: Statistical Process Control, SPC, chemical solutions

Statistical Process Control (SPC) is one of the most effective methods used to achieve and assure quality parts and products. The main tool used in SPC is the control chart, which monitors process performance over time. Traditional control chart methodology assumes that process output tends to vary randomly over time around some average value. Any drifts or trends are considered unnatural to the process and will be flagged as such by the control chart. However, trends over time are likely to be natural and expected features in some chemical solutions. To avoid flagging these inherent features as unnatural to the solution, the traditional control chart methodology has been modified. This paper describes the modified methodology and how it is being applied to monitor key components in plating and cleaning solutions at Pratt & Whitney Rocketdyne in West Palm Beach, FL.

Aliasing Strings for Some Nonregular Designs with Generalized Minimum Aberration

❖Debra Ingram, Arkansas State University, Department of Mathematics and Statistics, PO Box 70, State University, AR 72467, *dingram@astate.edu*; Alison L. Wiley, Arkansas State University

Key Words: fractional factorial design, generalized resolution, Hadamard matrix, nonregular design, partial aliasing, projection

Nonregular designs of 20, 24, and 28 runs help fill the gap in available run sizes between the 16-run and the 32-run regular 2^(k-p) fractional factorials, and provide experimenters flexibility in run size. For several values of k, the nonregular designs offer higher generalized resolution than their regular counterparts. These nonregular designs exhibit better projection properties and may provide advantages in satiability of effects through partial (instead of full) aliasing. The aliasing structures for the regular 2^(k-p) designs are straightforward and can be summarized by a defining relation. The aliasing structures for nonregular designs, on the other hand, are more complex. This paper provides the aliasing strings for some nonregular designs, specifically those that are viewed as attractive alternatives to the 16-run and 32-run regular designs.

Semiparametric Maximum Likelihood for Regression Models with Missing Data

❖Suzanne Dubnicka, Kansas State University, Department of Statistics, 101 Dickens Hall, Manhattan, KS 66506, *dubnicka@ksu.edu*

Key Words: Horvitz-Thompson estimator, kernel density estimation, maximum likelihood, missing at random

Data sets from larger studies, such as AIDS Clinical Trials Group protocol 175 (ACTG 175), are often incomplete. Two methods which allow for the analysis of incomplete data are the EM algorithm and weighted estimating equations (WEE). We study the problem of estimating the parameters of a regression model when the error distribution is completely unknown and either the response variable or some of the covariates are missing at random. The proposed estimation procedure replaces the true unknown error density in the likelihood with a kernel density estimate and maximizes the estimated likelihood. The kernel density estimate is based on the Horvitz-Thompson estimator which also forms the basis for WEE. Thus, the proposed methodology shares common elements with both the EM algorithm and WEE. The method is illustrated on the ACTG 175 data and is verified via Monte Carlo simulation.

Portmanteau Test of Independence for Functional Observations

❖Robertas Gabrys, Utah State University, Department of Mathematics and Statistics, 3900 Old Main Hill, Logan, UT 84322, *gabrys.robertas@gmail.com*; Piort Kokoszka, Utah State University

Key Words: Functional Observations, Functional Principal Component Analysis, Independence Test

We proposed a chi-squared test for independence and identical distribution of functional observations. Asymptotic theory based on functional principal components expansions and Hilbert space techniques is developed. In a simulation study, the test is shown to have good empirical size and power. Its application is illustrated on two data sets: credit card sales activity and geomagnetic records.

Generalized Inverse Characteristics in Linear Statistical Modeling

❖Fang Duan, Sam Houston State University, 2106 Avenue J Apt F, Huntsville, TX 77340, *duanfang@gmail.com*; Cecil Hallum, Sam Houston State University

Key Words: generalized inverses, linear models, model constraints

Four key categories of generalized inverses and their applications in statistics, especially in general linear models are investigated in this study. The investigation will include coverage of the various general inverses referred to as 1) the generalized inverse, 2) the reflexive generalized inverse, 3) the normalized generalized inverse, and 4) the pseudoinverse. The various generalized inverse properties will be investigated in regard to their applications in statistics. These properties will be discussed through their applications particularly in the restricted general linear model , whereby the parameter is constrained further to satisfy the equality constraint . Also investigated is their contribution to solutions of systems of equations, in ridge regression, analysis of variance, and analysis of covariance models. Other peculiar properties will be highlighted as well.

Testing for Lack of Dependence in a Functional Linear Model

❖ Inga Maslova, Utah State University, 3900 Old Main Hill, Department of Math and Stat, Logan, UT 84322-3900, *inga@cc.usu.edu*; Piort Kokoszka, Utah State University

Key Words: functional, linear, model, test of independence

We study a functional linear model, $Y = \Psi X + \epsilon$ with functional response and explanatory variables; Ψ is a bounded operator in the Hilbert space L^2. We propose a simple test of the nullity of Ψ based on the principal component decomposition. The test statistic has asymptotic chi-square distribution, which is also an excellent approximation in finite samples. The methodology is applied to magnetometer data. Records from twelve magnetic observatories during disturbed and quiet periods are analyzed to check if there is functional linear dependence between observations at different latitude stations.

A Comparison of Restricted Heteroscedastic Estimators for Small Samples (n<50)

❖ Luis Frank, Universidad de Buenos Aires, Sucre 2020 P12A, Buenos Aires, C1428DVJ Argentina, *lfrank@mecon.gov.ar*

Key Words: regression analysis, heteroskedasticity, White estimator, OLS, GLS, efficient estimator

Consider the model y=Xfl+e subject to Rfl=r, such that e~N(0,V) - V diagonal of unequal elements - and X is a matrix of constants of size nxk. Both E(fl)=fl* and cov(fl*) - fl restricted - may be drawn from Akaike's or White's solutions for non-restricted models introducing an estimated V* from OLS residuals. Although these solutions converge to efficient estimators of fl* when n goes to infinity, there is little evidence about their performance for small sample sizes (n< 50). The paper compares six estimators of fl* and computes by simulation the sum of squares s=(fl*-fl)'(fl*-fl) for n={5, 10, 15, 20, 25, 50} and k=4. The findings suggest that OLS is the most efficient estimator for extremely small samples (5< n< 10), whereas White's estimator corrected by Davidson-Mac Kinnon's factor is the most efficient for small- to medium-sized samples (25< n< 50). Results are ambiguous for 10< n< 25.

Methods for Constructing Classifiers from Sparse Categorical Data for Forensic Document Analysis

❖ Christopher Saunders, George Mason University, 4400 University Drive Ms1G8, Fairfax, VA 22030, *csaunde6@gmu.edu*; Linda J. Davis, George Mason University

Key Words: Categorical Data, Forensic Document Analysis, Pattern Recognition, Bayes classifiers, Biometric identification

Recent challenges to the admissibility of handwriting evidence have illustrated the need for a statistical approach to the problem of writer identification. Discriminant functions are constructed with a dataset containing one hundred distinct writers who have each submitted five script copies of the London paragraph. Each character in the paragraph is represented as a graphical isomorphism; reducing each paragraph to the frequency of graphs used to write each character. This presentation compares approaches to constructing plug-in Bayes classifiers based on the categorical data that predict the writer-ship of an unknown document. The classifiers are compared using leave-one-out cross-validation.

Testing Equality of Mean Functions for Continuous Time Stochastic Processes

❖ Yolanda Munoz, University of Texas at Houston, 1200 Herman Pressler, RAS w834, Houston, TX 77030, *yolanda.m.munoz@uth.tmc.edu*

Key Words: u-statistic, non-parametric methods, jackknife, funtional data analysis, ANOVA

A test statistic is developed to determine whether two or more groups of functions, derived from continuous time stochastic processes, belong to the same population. An approximate chi-square test is proposed for comparing n group means. The test statistic considered in this paper uses a quadrature approximation to the L_2 norm. The proposed test statistic is shown to be a U-statistic and its asymptotic properties are formulated. A small Monte Carlo study is conducted and we illustrate its application with a real data set.

Analysis of Within-Individual Areas Under the Curve Weighted by an Individual Variance

❖ Jeff Thostenson, University of Arkansas for Medical Sciences, 4301 W Markham Street, Slot 781, Little Rock, AR 72211, *jdthostenson@uams.edu*; Reid Landes, University of Arkansas for Medical Sciences

Key Words: Trapezoidal rule, Simpson's rule, Weighting, Repeated measures

The area under a curve (AUC) can serve as a powerful summary measure in diverse disciplines. Often, analyses of AUC data assume equal variance across individuals, which may not be true. Weighting each individual's AUC with its inverse variance can sharpen inference. A novel method for calculating the variance of an individual's AUC is provided and implemented on two types of AUC data: monotone decreasing (increasing) and nonmonotonic curves. From botany a study of tree bud break, and from psychology a study of impulsivity are analyzed as illustrative examples.

Spatio-Temporal Modeling Using Stochastic Differential Equations

❖ Gavino Puggioni, Duke University, 222 Old Chemistry Bldg, Statistics, Box 90251, Durham, NC 27708-0251, *gavino@stat.duke.edu*

Key Words: Bayesian, SDE, Diffusion, Soil Moisture, Nonlinear

In this work we present some strategies to efficiently estimate a system of stochastic differential equations that are spatially correlated. Our approach is based on a data augmentation procedure that allows a more precise approximation of the likelihood. An application to soil moisture dynamics is provided.

Spatial Statistics for Entrepreneurial Opportunities

❖ Jerry Platt, University of Redlands, 1200 E Colton Avenue, Redlands, CA 92373, *jerry_platt@redlands.edu*

Key Words: GIS, Spatial, Business

Recently, the University of Redlands received a $1 million grant from the U.S. Small Business Administration (SBA). The grant program is a collection of initiatives that focus on spatial solutions to small business problems and opportunities. The three primary components of the applied research component of the grant are regional market analysis, geo-referenced functional research, and GIS case studies. All three components have a spatial statistics perspective. This presentation summarizes grant methodology and findings, with an emphasis on linking statistical analyses to business maps of customers, competitors, and the resultant entrepreneurial opportunities. Open-source software for statistical modeling and spatial analysis are emphasized.

A Bayesian Model and Expansion for Predicting Undergraduate and Graduate Institutional Re-Enrollment Rates and Totals

❖Lindsay Renfro, Baylor University, Department of Statistical Science, One Bear Place #97140, Waco, TX 76798, *lindsay_renfro@baylor.edu*; Mark Hamner, Texas Woman's University

Key Words: Bayesian inference, re-enrollment totals, enrollment prediction, re-enrollment rates

A university's student enrollment and the revenue it generates for the institution are key for future financial and infrastructure planning. In particular, having prior knowledge of expected re-enrollment in an upcoming semester is crucial, since re-enrolled students often comprise the majority of students enrolled in any semester. Thus, we employ a Bayesian model to predict re-enrollment rates and totals in a given semester of interest for students where a priori data exist. Here, we derive information through innovative partitioning of Texas Woman's University student data and use SAS to expose consistent re-enrollment patterns. From these patterns, only a few variables common to most institutions of higher education and only one year's worth of data are necessary for construction of an efficient Bayesian model and an expansion of its accompanying prediction.

Uses of Packages in R and S-PLUS

❖Stephen Kaluzny, Insightful Corporation, 1700 Westlake N 500, Seattle, WA 98109, *spk@insightful.com*

Key Words: packages, R, S-PLUS, software, textbooks

Both the R and S-PLUS implementations of the S language now have a package system for extending the system. An S package can include S language code, C/Fortran code, data and documentation for the package. While package system is used primarily to add new computational techniques to the language, it can be used to share data, analysis methods and documentation in project collaborations. Packages can also be used to distribute example data sets as well as analysis scripts for classes and textbooks. This poster will give an overview of the key features of the package system and show examples of the various uses of packages.

Manipulatives To Improve Learning DOE Concepts

❖James Alloway, EMSQ Associates, 152 Grandview Ave, Catskill, NY 12414, *jalloway@earthlink.net*

Key Words: DOE, Teaching Statistics, Geometric Models, Hands-on Learning

Statistically designed experiments are powerful tools for process improvement and optimization, yet are underutilized because new students don't feel comfortable with the methodology. Typical introductory courses present columns of positive and negative signs to show the design levels, then jump directly to main effects plots and interaction plots to summarize results. The concepts of how the design is created and the physical meaning of effects and interaction are skipped. A unique manipulative for two level factorials allows students to visualize the design space as they plan the design, insert response data into the manipulative as the experiment is run, visualize the nature and magnitude of the effects in the model, and finally calculate values for the effects. This approach clarifies the concepts of DOE, making it more likely that these techniques will be utilized.

The Progress of Master of Applied Statistics Program at Penn State

❖Mosuk Chow, The Pennsylvania State University, 315 Thomas Building, University Park, PA 16802, *mchow@stat.psu.edu*

Key Words: Master of Applied Statistics, Professional Science Master, Internship, Business, Industry and Government

The Master of Applied Statistics Program at Penn State is a Professional Science Master program. Support for the development of this program has been provided by the Alfred P. Sloan Foundation in 2001-2004. The program aims to provide its graduates with broad knowledge in a wide range of statistical application areas with employable, highly sought after skills Statistics demanded by government agencies, consulting firms and industries. The target program size of the program is to enroll 15 students per year. There is a trend of an increasing application pool to our program. Over the years, we have attracted high quality students with a diversified backgrounds. From our experience, internship opportunities are an invaluable experience for the M.A.S. students and we have strong interest in continuing to develop internship relationship with companies and government agencies.

Passing On the Passion: Recruiting Students of Statistics

❖Reid Landes, University of Arkansas for Medical Sciences, 4301 W Markham St, Slot 781, Little Rock, AR 72205-7199, *rdlandes@uams.edu*

Key Words: career choice, statistical education, undergraduates

It's pretty simple: we're statisticians because we love statistics. And we believe that if we could get students to really listen to us, they might just end up falling in love with statistics, too. But why aren't they listening to our song? How can we get them to tune in, and then join us in our dance? Reviewing the literature, we pick out the sour notes of recruitment and highlight the motifs of recruiting strategies employed within statistics and among other disciplines. We tune strategies that have gotten off key, and provide an activity for churning out hit after recruiting hit.

WeBWorK in the Statistics Curriculum

❖Jeff Holt, University of Virginia, Department of Statistics, PO Box 400135, Charlottesville, VA 22904, *jjholt@virginia.edu*

Key Words: webwork, computer, homework

This poster will describe features of the WeBWorK online homework system, currently being used in undergraduate statistics courses at various colleges and universities. This free system provides instant feedback, automatic grading, and is highly flexible. Information about assessment of courses using the system will be provided. If possible, a live demonstration of WeBWorK will be given.

Classroom Response Systems and the Development of Students' Conceptual Understanding: The Value of Question Ambiguity

❖Robert Terry, University of Oklahoma, 3140 Walnut Rd, Norman, OK 73072, *rterry@ou.edu*; Teri J. Murphy, University of Oklahoma; Curtis C. McKnight, University of Oklahoma; Michael Richman, University of Oklahoma

Key Words: Psychometrics, Item-Writing, Clickers, Validity

Classroom response system (CRS) technology, used successfully in introductory physics courses for a number of years, can facilitate student-student or whole-class discussion of statistics concepts. One challenge in using such teaching strategies is determining what constitutes a useful question. The CRS technology, consistent with standard psychometric theory, assumes that one of the MC options is clearly better than the distracters, assuming that instructors will give students full credit for choosing this option and

less credit for choosing one of the distracters. As we have used this technology, however, we have discovered that impressive discussion has been generated particularly by questions for which more than one response can be defended. Such questions include the choice of appropriate population, the choice among different estimators, and the handling of outliers.

Comparing Two Proportions with the Binomial Distribution

❖James B. Orris, Butler University, 140 Holcomb Research Bldg., 4600 Sunset Ave, Indianapolis, IN 46208, *orris@butler.edu*

Key Words: hypothesis testing, binomial distribution, p-value, two proportion test

The hypothesis testing module in most introductory statistics courses shows how to use the normal distribution to test a sample proportion vs. a hypothesized value. Some courses/textbooks also show how to get the exact p-value using the binomial distribution. This paper shows how students can extend this technique to compare two proportions by calculating a binomial distribution for each proportion and then calculating the joint probabilities from which the p-value can be determined. The method discussed in this paper is useful as a teaching technique because it reviews probability distributions and probability concepts. The calculations will be illustrated with an Excel spreadsheet that makes a good project for students. An Excel VBA program will also be presented that does the calculations and compares the results to the traditional normal curve method.

Expectations for Statistical Literacy: A Survey of College Faculty

❖S. David Kriska, Restat Systems, 3758 Surrey Hill Place, Upper Arlington, OH 43220, *davidkriska@sbcglobal.net*; Mark C. Fulcomer, Richard Stockton College of New Jersey; Marcia M. Sass, University of Medicine and Dentistry of New Jersey

Key Words: Statistical literacy, introductory statistics, statistical tasks

Statisticians will certainly agree on the need for statistical literacy among college graduates, but specifying what constitutes statistical literacy is debatable. As part of an effort to better understand the concept, Kriska, Fulcomer, and Sass (2006) proposed the development of a 68 item survey instrument applicable to a variety of situations. This presentation describes the use of the resulting survey instrument with university faculty responding from perspectives of the literacy demands for a range of specific courses they teach. The results include an overall item analysis summarized over all courses included in the sample as well as comments on specific findings by various areas.

Integrating Video Clips into an Introductory Statistics Course

❖Edward Mansfield, The University of Alabama, ISM, Box 870226, Tuscaloosa, AL 35487-0226, *emansfie@cba.ua.edu*; Subhabrata Chakraborti, The University of Alabama

Key Words: Teaching, Learning, Applications

The Against All Odds video series is a collection of 26 half-hour episodes depicting uses of Statistics in real situations. Part of the Annenberg/CPB Collection and produced by COMAP, the series has appeared on many public television stations. Although released in 1989, much of the material presented, although dated, still delivers a powerful punch. To facilitate use in the classroom a detailed timeline of each episode was presented at JSM-1995; segments highly effective for classroom learning were identified. However, the logistics of locating a desired segment on a videotape can be time consuming. This presentation offers several digitized short clips from the Against All Odds series edited into separate files which can be downloaded for easy playback in the classroom or for use as a supplement to homework assignments. Expose your students to the power of Statistical stories.

Classroom Simulation: Distributions of Sales and Queue Lengths at a Fast-Food Counter

❖Eric Suess, California State University, East Bay, Department of Statistics, 25800 Carlos Bee Blvd, Hayward, CA 94542, *eric.suess@csueastbay.edu*; Bruce Trumbo, California State University, East Bay; David Ahlberg, California State University, East Bay

Key Words: Simulation in R, random sum of random variables, queue length, exponential/uniform relationship, teaching

Lunch-hour sales at a fast food outlet are modeled and simulated using R. Concepts illustrated include: expectation, variance, and distribution of the sum of a random number of random variables; some relationships among binomial, Poisson, uniform, and exponential distributions; and average length of an M/M/s queue. Level suitable for upper-division probability students.

Using JMP Scripts in Teaching Introductory Statistics

❖Amy Froelich, Iowa State University, 324 Snedecor Hall, Department of Statistics, Ames, IA 50011-1210, *amyf@iastate.edu*; William M. Duckworth, Creighton University

Key Words: simulation, classroom activities, applets

Simulation has become an important tool in teaching topics related to sampling distributions and inference in the introductory statistics class. Many of these simulations have been developed with Java applets and made available on the web. While these applets are easy to use and readily available to statistics instructors, they may not match classroom, laboratory or homework activities. As a result, students can struggle with the transition between the classroom and the computer simulation activities. We will show how scripts in JMP can be used to develop simulation activities like the ones found on the web that match classroom, laboratory or homework activities.

The International Statistical Literacy Project (ISLP) of the International Association for Statistics Education (IASE): How To Use It for Teaching and Learning

❖Juana Sanchez, University of California, Los Angeles, International Statistical Literacy Project of IASE, 8125 Math Sciences Blg Box 951554, Los Angeles, CA 90095, *jsanchez@stat.ucla.edu*

Key Words: Statistical literacy, ISLP, IASE, teaching and learning, National Statistical Offices, assessment

The International Statistical Literacy Project (ISLP) is an ongoing project of the International Association for Statistics Education (IASE) with the objective of promoting statistical literacy around the world and aiding those interested in statistical literacy. Its web page at www.stat.auckland.ac.nz/~iase/islp contains the largest collection of international statistical literacy resources available today relevant for the media, teachers, students, adult educators, national statistical offices, and many others. In this poster, we illustrate how to make the most of these resources with several examples on how to use them: assessment tools built from it, community programs proposals to improve statistical literacy, national statistics tutorials and many other. In each of these examples, we show in detail how we put it together using the ISLP.

Applying Some Computational Simplifications for Binomial and Poisson Random Variables to the Teaching of Statistical Concepts

❖Mark C. Fulcomer, Richard Stockton College of New Jersey, 48 Trainor Circle, Bordentown, NJ 08505, *mcfulcomer@aol.com*; S. David Kriska, Restat Systems; Maritza Jauregui, The Richard Stockton College of New Jersey; Marcia M. Sass, University of Medicine and Dentistry of New Jersey; Jallah M. Kennedy, University of Medicine and Dentistry of New Jersey

Key Words: teaching statistics, binomial distributions, Poisson distributions, computational formulae

Among the simplest forms of random variables and with seemingly vast potentials for "real" applications, binomial and Poisson distributions are discussed in many introductory textbooks on probability and statistical methods. However, despite initial appearances, the uses of both distributions in actual practice are often limited by computational difficulties as N increases. This presentation describes the incorporation of simplifying "recursive" formulae into two Excel spreadsheet programs that combine graphics with the ability to handle "large" problems. Utilizing textbook and other health-related examples in classroom settings, these programs have enhanced illustrations of several important and often challenging statistical concepts, including sampling distributions, hypothesis testing, the Central Limit Theorem, and normal approximations to both distributions.

A Free Software Live DVD for Statistics

❖ Panayotis Giannakouros, University of Missouri-Kansas City, 2139 Collinway 4, Ottawa Hills, OH 43606, *lists@potis.org*; Lihua Chen, The University of Toledo

Key Words: R, free software, open source

Configuring and teaching a system of interacting software packages can be problematic for an instructor. Even if some computer labs can be properly set up, users may want to work at a variety of campus and noncampus locations, making the investment in learning a set of tools difficult to motivate. We describe a suite of free and open source software packages relevant to statistics and demonstrate how these packages can be configured to integrate with each other and to run from a DVD on a Mac or PC. We find the DVD-burning capability available on newer desktop computers makes the distribution of a complete statistical environment feasible. This rapidly deployable and self-contained format is well suited to trying a new environment and to extending flexibility to users. We also report on student acceptance and use of these tools at the University of Toledo.

A New Type of Portmanteau Test Based on the Discrete Cosine Transform

❖ Sungun Oh, Sookmyung Women's University, Department of Statistics, Chungpa-dong Yongsan-gu, Seoul, 140-742 South Korea, *ruma0819@hanmail.net*; Hyemin Cho, Sookmyung Women's University; In-Kwon Yeo, Sookmyung Women's University

Key Words: Discrete cosine transform, Multivariate portmanteau test, Ljung-Box test, Determinant of covariance

We present a new type of portmanteau test in the frequency domain which is derived from the discrete cosine transform (DCT). For the stationary time series, DCT coefficients are asymptotically independent and their variances are expressed by linear combinations of autocovariances. The covariance matrix of DCT coefficients for white noises is diagonal matrix whose diagonal elements is the variance of time series. A simple way to test the independence of time series is that we divide DCT coefficients into two or more sets and then compare their sample variances. We also do this by testing the slope in the linear regression model of which the response variables are absolute values or squares of coefficients. Similar approaches are applied to multivariate time series. Simulation results show that the proposed tests have much higher powers than existing tests in most cases of our experiments.

Prediction Interval Estimation in Transformed ARMA Models

❖ Hyemin Cho, Sookmyung Women's University, Chungpa-don, Yongsan-gu, Seoul, 140-742 South Korea, *coolmin98@hanmail.net*; Sungun Oh, Sookmyung Women's University; In-Kwon Yeo, Sookmyung Women's University

Key Words: ARMA models, Coverage probability, Prediction interval, Smearing estimation, Yeo-Johnson transformation

One of main aspects of time series analysis is to forecast future values of series based on observations up to a given time. The prediction interval for future values is usually obtained under the normality assumption. When the assumption is seriously violated, a transformation of data may permit the valid use of the normal theory. We investigate the prediction problem for future values in the original scale when transformations are applied in ARMA models. In this study, Yeo-Johnson transformation which can solve the skewness of data is applied and the smearing estimation is employed to reduce the bias. We also estimate the prediction intervals. We present conditions that ensure that the intervals have asymptotically correct coverage probability. Simulations show that the coverage probabilities of proposed intervals are closer to the nominal level than those of usual intervals.

295 Business and Economics Statistics Section Speaker with Lunch (fee event)

Business and Economics Statistics Section

Tuesday, July 31, 12:30 pm–1:50 pm

The Use of Birth Panel Data in Economic Research

❖ Jason Abrevaya, Purdue University, Department of Economics, 403 West State Street, West Lafayette, IN 47907-2056, *abrevaya@purdue.edu*

The use of maternally linked birth data (at the federal and state levels) for economic research will be discussed. Specific applications will be discussed, including the effects of public policy (e.g., cigarette taxes) upon prenatal behavior and the effects of child gender mix upon fertility outcomes.

296 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section

Tuesday, July 31, 12:30 pm–1:50 pm

Joining the Evil Empire: What It's Like To Move from Academics to Industry

❖ Robert Parker, Amgen Inc., 1 Amgen Center Drive, Mail Stop 24 2 C, Thousand Oaks, CA 91320, *raparker@amgen.com*

Key Words: Academics, Evil, Industry

For all the academics out there secretly wondering "what's it like in industry," this seminar will answer your questions. Secrets will be revealed!

Learn some of the good–and bad–of industry from someone who recently changed sides.

Statistical Analysis Plans: How, When, Why, and for Whom Do We Do Them?

❖Bruce Binkowitz, Merck & Co., Inc., 18 Bates Way, Westfield, NJ 07090, *binkowitz@merck.com*

Key Words: ICH E9, analysis, plan

Statistical analysis plans (SAPs) are suggested by guidance (e.g., ICH E9) as appropriate documents for prespecifying statistical methodology and data handling rules. These analysis plans should be completed prior to the unblinding of a database. This roundtable is intended to discuss why we write these documents, who the target audiences and reviewers are, when in the lifecycle of a trial do we write these, are they really necessary, and how easy/difficult they are to put together.

JAMA Policy on Independent Analysis and Public Trust

❖Hong Liu-Seifert, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *liu-seifert_hong@lilly.com*

Key Words: JAMA, independent analysis, independent statistician, transparency, public trust

JAMA's policy requires independent analysis by independent stats for industry-sponsored papers. Join us for an open and productive discussion on approaches to address similar policies and transparency issues and to improve public trust in our industry as a whole.

Blinded Sample Size Re-estimation

❖Biao Xing, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, *xing.biao@gene.com*

Key Words: Sample Size, Power, Adaptive Design

Blinded sample size re-estimation allows for modifying the sample size of an ongoing trial to ensure sufficient statistical power without breaking the blind. One challenge is the blinded estimation of the within group variance. Early proposed methods either make untenable assumptions or are only applicable to two-treatment trials. Moreover, these methods are often biased. We propose a simple, unbiased method, that also makes minimal assumptions. The method uses the enrollment order of subjects and the randomization block size to estimate the variance and then re-estimate the sample size. It can be applied to normal or non-normal data, to trials with two or more arms, equal or unequal allocation schemes, and fixed or random randomization block sizes. Results from simulations and data analysis suggest the proposed blinded sample size estimation approach is practical.

Estimation of Treatment Effects in Randomized Trials with Noncompliance and Crossover: Why ITT Analysis May Not Be Our Best Option

❖Afisi S. Ismaila, McMaster University, Clinical Epidemiology and Biostatistics, 1280 Main St. West, T-27 Building, Room 324, Hamilton, ON L8S 4L8 Canada, *ismailas@mcmaster.ca*; Stephen Walter, McMaster University

Key Words: Randomized controlled trials, compliance, efficacy, treatment preference

When there are departures from assigned treatments in a randomized controlled trial (RCT), researchers are faced with the problem of estimating treatment benefits in compliant and crossover patients while respecting randomization. We will discuss why some of the well-known traditional approaches may not be our best option. We also will discuss some of the newest approaches for dealing with this problem.

297 Section on Bayesian Statistical Science Roundtable with Lunch (fee event)

Section on Bayesian Statistical Science

Tuesday, July 31, 12:30 pm–1:50 pm

Meta-Analysis of Diagnostic Tests

❖Vanja Dukic, The University of Chicago, 5841 S Maryland Ave, MC 2007, Chicago, IL 60637, *vanja@uchicago.edu*

Key Words: meta-analysis, diagnostic test, bayesian models, software, decision making, evidence synthesis

The purpose of this roundtable is to gather researchers interested in meta-analysis and facilitate a discussion among them about recent statistical methods for evidence synthesis, with the focus on medical decisionmaking. Particular attention will be given to methods for Bayesian meta-analysis of diagnostic test studies. Meta-analysis software also will be discussed.

298 Section on Government Statistics Roundtables with Lunch (fee event)

Section on Government Statistics

Tuesday, July 31, 12:30 pm–1:50 pm

Data User Issues with the American Community Survey

❖Patricia Becker, APB Associates/APDU/SEMCC, 28300 Franklin Rd, Southfield, MI 48034, *pbecker@umich.edu*

Key Words: ACS, Census, Data users

Almost a year ago, we saw the release of the first full year of data collection, the 2005 ACS. Data were released in several formats on American FactFinder. What are user experiences with these products? How can they be improved? What should we expect the U.S. Census Bureau to do so the products work better for us?

Remote Access to Microdata

❖Julia Lane, NORC at the University of Chicago, 55 E Monroe St, Suite 4800, Chicago, IL 60603, *gasparac-julie@norc.org*

Key Words: data access, confidentiality, remote access, microdata

This roundtable will focus on discussing the costs and benefits of providing researchers with remote access to microdata. It will describe the implementation at NORC using NIST's Advanced Technology Program data. It also will describe some of the metrics being developed to measure the impact of the approach on the quality of analysis and the associated risk.

299 Section on Health Policy Statistics Roundtables with Lunch (fee event)

Section on Health Policy Statistics

Tuesday, July 31, 12:30 pm–1:50 pm

Recent Methodological Developments on Prediction of Health-Care Costs

❖ Xiao-Hua (Andrew) Zhou, University of Washington, 1705 Pacific Street NE, Seattle, WA 98195-7232, *azhou@u.washington.edu*

Key Words: prediction, health care, costs

In this roundtable, I will discuss the issue and difficulty of predicting future health care costs of individual subjects due to their special distributional features, including skewness, zero, and heteroscedasticity. Then, I will discuss recent attempts to overcome these difficulties, including parametric, semiparametric, and nonparametric models, and contrast their strengths and limitations. If time allows, I will illustrate these methods in a real clinical study, conducted in the population of U.S. veterans.

Statistical Methods for Measuring Workflow and Workload in Medical Care Systems

❖ Jimmy Efird, John A. Burns School of Medicine, 651 Ilalo Street BSB320B, Honolulu, HI 96813, *jimmy.efird@stanfordalumni.org*

Key Words: Information systems, medical system, workflow & workload analysis, hospitals, time costs

Millions of dollars have been invested in complex information technology systems to improve patient outcomes, enhance bedside care/patient monitoring and minimize drug-dispensing errors throughout the medical care system in the United States. Unfortunately, rigorous statistical methods are not used to ascertain the impact these new technologies may have on the workflow and workload of medical staff, which involves changes in their time and task commitments following related changes in the work environment. Often, these new systems are removed from clinic floors, following numerous complaints from doctors and nursing staff. This luncheon will focus on the statistical methodology for conducting workflow and workload analysis. Topics discussed will include measures of accuracy and reliability, sampling methodology, categorical responses, and repeated measures designs.

300 Section on Physical and Engineering Sciences Roundtables with Lunch (fee event)

Section on Physical and Engineering Sciences

Tuesday, July 31, 12:30 pm–1:50 pm

Challenges in Mixture Experiments

❖ Greg Piepel, Battelle/PNNL, PO Box 999, Richland, WA 99352, *greg.piepel@pnl.gov*

Key Words: Mixture experiment, Experimental design, Data analysis, New methods, Roundtable

A mixture experiment (ME) involves mixing components to form an end product and measuring a response, which depends on the proportions of the components and maybe other factors (e.g., total amount, process variables). Methods for ME design and analysis have been developed over 50 years. However, challenges often exist for ME problems. One challenge is to decide which of several approaches is best for a given ME problem. Another challenge might be a "twist" in the ME or its data analysis not addressed by existing methods or software. Roundtable participants should come prepared to briefly describe one or two challenges they have faced in ME problems, what they did to address the challenges, and whether new methods or software capabilities are needed. The opportunity for discussion and feedback from other participants should make this an interesting and useful roundtable.

Drug Discovery and Cheminformatics: Advances, Challenges, and Frontiers

Raymond L.H. Lam, Centocor, Inc.; ❖ Jacqueline M. Hughes-Oliver, North Carolina State University, Department of Statistics, Room 203C Patterson Hall, 2501 Founders Drive, Raleigh, NC 27695-8203, *hughesol@stat.ncsu.edu*

Key Words: Cheminformatics, High Throughput Screening, Drug Discovery

Six exploratory centers for cheminformatics research were established in 2005. These centers serve as precursors to full-scale cheminformatics research centers for the development of computational algorithms, software tools, theoretical insights, and data-rich resources applicable to chemical diversity and high-throughput screening. A major challenge issued in the NIH request for application was to create cross-disciplinary teams to bring new insight for addressing the difficult problems within cheminformatics. Statisticians have been taking a leading role and making major contributions toward this effort. This roundtable discusses advances, challenges, and frontiers in cheminformatics research. Folks from other areas who manage or are involved in similar cross-disciplinary efforts also are welcome.

301 Section on Quality and Productivity Roundtables with Lunch (fee event)

Section on Quality and Productivity

Tuesday, July 31, 12:30 pm–1:50 pm

Tools for Planning and Properly Sizing a Factorial DOE

❖ Shari Kraber, Stat-Ease, Inc., 2021 E Hennepin Ave, Ste 480, Minneapolis, MN 55413, *shari@statease.com*

Key Words: DOE, Planning Process, Power, factorial design

Newcomers to factorial design find it difficult to choose appropriate designs with adequate power. During this luncheon, we will discuss a clear process to determine the best design that fits the problem. Attendees are welcome to bring their planning processes to share with the group. A discussion of statistical power will demonstrate how the size of the effect relative to the noise is a critical criterion in design selection. We also will discuss how to choose ranges for the input factors, the importance of evaluating aliases, and checking runs for safety. Various case studies will illustrate the importance of using a planning process to avoid DOE failure due to an incorrect design choice. Attendees will take away a variety of strategies for determining which factorial design is appropriate for their data analysis needs.

Teaching Experimental Design to Students Planning an Industrial Career

❖John Lawson, Brigham Young University, 223B TMCB, Provo, UT 84602, *lawson@byu.edu*

Key Words: Experimental Design, Statistical Software

We will discuss topic emphasis in experimental design classes taught for students planning a career as an industrial statistical consultant. Software has made both the construction of experimental designs and analysis of the resulting data much easier. Therefore, students planning industrial careers have less need to focus on the hand computational formulas or algorithms to construct certain types of designs that are presented in many traditional experimental design texts. More time could be focused on teaching students, a strategy for selecting a design that is appropriate for the research objective, how to recognize the appropriate error term for hypothesis testing, and how to interpret and present results that will be meaningful to nonstatisticians. Participants should be prepared to discuss their experiences and share what they have found to be effective.

302 Section on Statistical Consulting Roundtables with Lunch (fee event)

Section on Statistical Consulting
Tuesday, July 31, 12:30 pm–1:50 pm

Alternative Models for Consulting Services in Academic Medical Centers

❖Shelley Hurwitz, Harvard Medical School, 1620 Tremont Street, BWH CCI 3, Boston, MA 02120, *hurwitz@hms.harvard.edu*

Key Words: consulting, academic medical center

As acknowledgment of the value of statistics in medical research expands, the need for statistical services grows. This will be an informative discussion for statisticians involved in consulting services at academic medical centers and for those considering such involvement. The organizer designed and directed two consulting services at an academic center associated with Harvard Medical School, with different operating models and fee structures. One had a large number of independent consultants over six years. The other had a smaller number of statisticians on staff for three years so far. Participants involved in consulting services will have the opportunity to describe their experiences. Attendees will learn what works and what does not and benefit from the experiences of the group in creating and/or participating in consulting services at academic medical centers.

Statistical Consulting in Medical Devices: Opportunities and Challenges

❖Jianxiong Chu, Food and Drug Administration, 1350 Piccard Drive, First floor, Rockville, MD 20850, *jianxiong.chu@fda.hhs.gov*

Key Words: statistical consulting, medical devices, opportunities, challenges

Compared to drugs, medical devices are as vital to prevent, diagnose, and treat diseases. The device industry has developed life-saving and life-enhancing products such as pacemakers, artificial joints/hips/spinal discs, and drug-eluting stents. Recent scientific advances in bioengineering, biomaterials, genomics, and nanotechnology may accelerate the innovation of new medical devices for delivering high-quality health care. In this roundtable, I will briefly describe the differences between drugs and devices to kickoff the discussion about the opportunities and challenges of statistical consulting in the development of new medical devices. For those of you who have been involved in medical devices, please come and share. For those of you who have not been involved but want to, please also come and learn more about the exciting field of medical devices.

303 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education
Tuesday, July 31, 12:30 pm–1:50 pm

My Favorite Case Study

❖Laura Chihara, Carleton College, 1 North College Street, Northfield, MN 55057, *lchihara@carleton.edu*

Key Words: undergraduate, case-studies, discussions

Do you have a favorite case study from a scholarly journal, newspaper, or magazine that you use in an introductory statistics class? We invite you to share this case-study with others and join a discussion about using case studies effectively in a small classroom setting.

Group Work and How To Assess It

❖William Harkness, The Pennsylvania State University, 318 Thomas Bldg, University Park, PA 16802, *wlh@stat.psu.edu*

Key Words: Group Projects, Assesment of Projects, Final Grades

I love having students do group projects, but are they valid assessment measures? Studies have shown that collaborative group work results in improved student performance (and I believe this). However, should scores on group projects be used in assigning grades? What are the pros and cons of doing so? Let's discuss this!

What Is So Great About Service Learning (SL), and How Much Work Is It?

❖Amy Phelps, Duquesne University, Rockwell Hall, 600 Forbes Avenue, Pittsburgh, PA 15282, *phelpsa@duq.edu*

Key Words: Service learning, pedagogy, authentic assessments, student reflection

Support for SL activities in higher education is growing. SL projects fit with the GAISE guidelines, but how does one use SL, how much work is it, and what do students gain? Student projects may achieve goals of the GAISE guidelines, but may lose the social relevance importance of data in the real world. Anderson (1999) emphasizes when a "project comes from a local source, students' interest and ownership of the project is strong." SL incorporates reflection assignments to further strengthen student learning. SL projects do require more effort, but students benefit from seeing the instructor fully engaged, expectations are higher to meet client needs, and students rise to the occasion. The process begins early and unfolds throughout the term. Positive outcomes are realized in student development and service to a nonprofit in need of statistical support.

304 Section on Statistics in Epidemiology Roundtable with Lunch (fee event)

Section on Statistics in Epidemiology

Tuesday, July 31, 12:30 pm–1:50 pm

Managing the Biostatistics Consulting Service in an Academic Hospital

❖Dianne M. Finkelstein, Harvard/MGH, 50 Staniford Street, Suite 560, Boston, MA 02114, *dfinkelstein@partners.org*

Key Words: consulting, academic hospital

A biostatistics consulting service can be an opportunity for developing new collaborations in a research hospital. However, it also can present challenges in terms of balancing priorities and time. This roundtable will be a discussion of how best to handle these challenges. In addition, we will talk about choices, such as how to charge for the service, the training required, the role in educating the investigator, etc.

305 Social Statistics Section Roundtable with Lunch (fee event)

Social Statistics Section

Tuesday, July 31, 12:30 pm–1:50 pm

From Soft to Hard Science: Will Survey Research Make the Transition?

❖Robert Fay, U.S. Census Bureau, 4700 Silver Hill Rd, Washington, DC 20233, *robert.e.fay.iii@census.gov*

Key Words: research, foundations, cognitive psychology, linguistics, ethics, scientific communication

Apart from my “day job,” I have studied scientific communication for several years and begun describing myself as a “daytime statistician/nighttime scientist.” I’d like to informally share some of my reflections on the state of survey research that stem from my investigations. Both scientists and the public occasionally distinguish “soft” and “hard” science without firmly agreeing on how or where to draw this boundary. I suggest that disciplines and individual researchers can be divided linguistically by the discourse structure of their writing. In the last century, the “hard” sciences have shared common communication standards that are much less reliably observed in “soft” science. But, change is possible. I will point to disciplines where the soft to hard transition has occurred or is currently taking place and what survey research could learn from them.

306 ASA College Stat Bowl II

The ASA, IMS, ENAR, WNAR, SSC

Tuesday, July 31, 2:00 pm–3:50 pm

Round 2

❖Winners from Session 1

Round 2 will consist of two games. Six players will advance from Round 1 to Round 2, three players going to each game. The six players will be determined by the winners of the four games, plus the two non-winning players that score the highest. The winners of the two Round 2 games will face off in the Finals. Each player will receive team points depending upon how they finish (4 points for champion, 3 points for runner up, 2 points to semifinalists not advancing, 1 point to participants not advancing to round 2). Team accumulating the most points will be declared team champion. Ties for team totals are broken by the highest average score for all members of the teams in question.

307 Functional Data Analysis ✪

IMS, Biometrics Section, Section on Nonparametric Statistics

Tuesday, July 31, 2:00 pm–3:50 pm

A Smoothing Spline ANOVA Approach to Fit Functional Linear Models

❖Yuedong Wang, University of California, Santa Barbara, Department of Statistics and Applied Probability, Santa Barbara, CA 93106, *yuedong@pstat.ucsb.edu*

Key Words: functional data, functional linear model, smoothing spline, ANOVA decomposition

Ramsay and Silverman (1997) proposed functional linear models (FLM) where the functional data can be (i) the dependent variable, (ii) independent variable, or (iii) both. We cast these FLMs in the framework of general smoothing spline ANOVA decomposition. Using the Canadian weather data and R functions in the ASSIST package, we illustrate alternative approaches to fit and make inference. If time permits, we will introduce semiparametric nonlinear mixed effects models and illustrate how to use them to register curves.

Interpretable Functional Generalized Linear Models

❖Gareth James, University of Southern California, Bridge Hall, 401R, Marshall School of Business, Los Angeles, CA 90089-0809, *gareth@usc.edu*; Peter Radchenko, University of Southern California

Key Words: Generalized Linear Models, Functional Regression, Dantzig Selector

Regression models to relate a scalar, Y, to a functional predictor, X(t), are becoming increasingly common. Work in this area has concentrated on estimating a coefficient function, beta(t), with Y related to X through the integral of beta(t) multiplied by X(t). A significant issue in the functional domain is that of producing an interpretable beta curve and several papers have recently been written on methods to generate curves with simple structure. In this talk we propose an approach for extending one of these procedures to the functional generalized linear models paradigm. This allows one to produce interpretable beta curves for a much wider class of situations, including classification problems such as logistic regression. We demonstrate our method on different data sets. In addition, interesting nonasymptotic theoretical bounds on the estimation error are presented.

Iterative Functional Principal Component Analysis for Correlation Reduction

❖Fang Yao, University of Toronto, 100 St. George Street, Department of Statistics, Toronto, ON M5S 3G3 Canada, *fyao@utstat.toronto.edu; fyao@stat.colostate.edu*; C. M. Lee, Colorado State University

Key Words: Asymptotic, Functional data, Penalized spline regression, Principal components, Smoothing, Within-subject correlation

We propose an iterative estimation procedure for performing functional principal component analysis. The procedure aims at functional or longitudinal data where the repeated measurements from the same subject are correlated. For the handling of the within-subject correlation, we develop an iterative procedure which would gradually reduce the dependence amongst the repeated measurements made for the same subject. An increasingly popular smoothing approach, penalized spline regression, is used to represent the mean trend. This allows straightforward incorporation of covariates and simple implementation of inference procedures for coefficients. The resulting data after iteration are theoretically shown to be asymptotically independent, which suggests that the general theory of penalized spline regression developed for independent data can also be applied to functional data.

308 Bayesian Model Selection

IMS, Section on Bayesian Statistical Science

Tuesday, July 31, 2:00 pm–3:50 pm

Bayesian Model Selection and Levy Random Field Priors

❖ Merlise A. Clyde, Duke University, Box 90251, Durham, NC 27708, *clyde@stat.duke.edu*

Key Words: kernel regression, support vector machines

We consider the problem of nonparametric regression. One approach is to represent the function in a series expansion using a linear combination of basis functions. Overcomplete dictionaries provide a redundant collection of generating elements, however, coefficients in the expansion are no longer unique. Despite the non-uniqueness, this has the potential to lead to sparser representations by using fewer non-zero coefficients. We show that Levy random fields are ideally suited for construction of priors on functions using these overcomplete representations for the general nonparametric regression problem, and provide a natural limiting generalization of priors for the finite dimensional version of the regression problem. We compare this approach to other penalized likelihood methods, such as the LASSO, and discuss methods for posterior inference.

Fully Nonparametric Bayesian Ensemble Modeling

❖ Robert McCulloch, The University of Chicago Graduate School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637, *robert.mcculloch@chicagogsb.edu*; Edward I. George, University of Pennsylvania; Hugh Chipman, Acadia University

Key Words: data-mining, markov chain monte carlo, trees

Suppose we would like to learn the relationship between y and a high dimensional vector x based on a limited number of observations. In "BART: Bayesian Additive Regression Trees" (2006), Chipman, George and McCulloch develop a fully Bayesian approach for discovering and drawing inference about an unknown function f based only on assuming $y = f(x) + e$ with iid normal errors. In the spirit of "ensemble models," BART approximates f by a sum of many simple regression tree models, each of which are kept small with a strong regularization prior. In this work, we further extend the flexibility of the BART approach by relaxing the simple iid normal error specification and replacing it with a Dirichlet process model for the errors. Various specification and prior choices are explored. The costs as well as the benefits of this more flexible approach are illustrated.

Multiple Shrinkage Predictive Densities for Linear Regression Models

❖ Xinyi Xu, The Ohio State University, Department of Statistics, 1958 Neil Ave., Columbus, OH 43210-1247, *xinyi@stat.ohio-state.edu*; Edward I. George, University of Pennsylvania

Key Words: density prediction, model uncertainty, Bayesian model averaging, multiple shrinkage

In this talk, we consider the problem of estimating the predictive density of a future quantity of interest based on a linear regression model, where exists a large number of predictors but only some of them are potentially relevant. Bayesian Model Averaging (BMA) is a general and powerful tool in this context since it incorporates model uncertainty by averaging over competing models with different predictor sets. We use the BMA approach to construct a class of multiple shrinkage predictive densities that dominate the "noninformative" uniform prior Bayes procedure and therefore are minimax. We also show that these multiple shrinkage predictive densities adaptively shrink toward the model most favored by the data, and achieve minimal risks under different degrees of model sparsity.

309 Technometrics Invited Session: Topics in Information Technology ● ✪

Technometrics, Section on Statisticians in Defense and National Security, Section on Quality and Productivity

Tuesday, July 31, 2:00 pm–3:50 pm

Embracing Statistical Challenges in the Information Technology Age

❖ Bin Yu, University of California, Berkeley, Department of Statistics, 367 Evans Hall, Berkeley, CA 94720, *binyu@stat.berkeley.edu*

Key Words: cyberinfrasctruture, massive data, streaming data, data fusion, computation, database

This talk examines the role of statistics in the cyberinfrastructure proposed by the CS community to integrate computer technologies into the very fabric of science and beyond. Areas of science and text mining are reviewed to motivate the challenges that we face for us to join the cyberinfrastructure design and taking advantage of the parallelism trend in computing. They include developing Exploratory Data Analysis (EDA) for massive data sets, using non-traditional mathematical results for insights in high dimensions, and frameworks for streaming data analysis, data fusion, and taking into account computation, transmission and compression constraints. We conclude that, for the healthy existence of our field, computer technologies must be further integrated into statistics, and statistical thinking must be integrated into computer technologies or into the making of the cyberinfrastructure.

Statistical Principles in Image Modeling

❖Ying Nian Wu, University of California, Los Angeles, Department of Statistics, 8125 Math Sci Bldg, Los Angeles, CA 90095, *ywu@stat.ucla.edu*; Jinhui Li, University of California, Los Angeles; Ziqiang Liu, University of California, Los Angeles; Song-Chun Zhu, University of California, Los Angeles

Key Words: Sparse coding, Markov random fields, Meaningful alignment

Images of natural scenes contain rich variety of visual patterns. In order to learn and recognize these patterns from natural images, it is necessary to construct statistical models for these patterns. In this article, we review three statistical principles for modeling image patterns, namely, the sparse coding principle, the minimax entropy principle, and the meaningful alignment principle. We examine these three principles and their relationships in the context of modeling images as compositions of Gabor wavelets. We show that these three principles correspond to three regimes of composition patterns of Gabor wavelets, and these three regimes are connected by the change of scale or resolution.

A Statistical View of the Transient Signals That Support a Wireless Call

❖ Diane Lambert, Google, 76 9th Avenue, New York, NY 10011, *dlambert@google.com*; David A. James, Novartis Pharmaceuticals; Chuanhai Liu, Purdue University; A. Buvaneswari, Alcatel-Lucent; John M. Graybeal, Alcatel-Lucent; W. Michael MacDonald, Alcatel-Lucent

A wireless call requires the cooperation of a dynamic set of base stations and antennas with control of the call changing in response to changes in signal strength. The communication between the mobile and the base station that is needed to manage the call generates a huge amount of data, some of which is seen only by the mobile placing the call and some of which is seen by the base station that manages the call. This talk will describe some of the subtleties in the data and a statistical model that takes the time dependent, spatial, and multivariate nature of the many wireless signals used to maintain a call into account. Online estimation and model validation using data taken from a commercial wireless network will also be discussed.

310 A Sensible Way To Design a Longitudinal Study ● ✪

WNAR, Section on Statistical Consulting, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Biometrics Section, Section on Survey Research Methods, ENAR

Tuesday, July 31, 2:00 pm–3:50 pm

Sample Size Determination for Hierarchical Longitudinal Designs with Differential Attrition Rates

Anindya Roy, University of Maryland, Baltimore County; ❖ Dulal K. Bhaumik, University of Illinois at Chicago, 1601 W Taylor St, M/C 912, Chicago, IL 60612, *dbhaumik@psych.uic.edu*; Subhash Aryal, University of Illinois at Chicago; Robert D. Gibbons, University of Illinois at Chicago

Key Words: Cost analysis, Dropouts, Mixed-effects, Power analysis, Profile analysis, Three-level nested design

We consider the problem of sample size determination for three-level mixed-effects linear regression models for the analysis of clustered longitudinal data. Three-level designs are used in many areas, but in particular, multicenter randomized longitudinal clinical trials in medical or health-related research. In this case, level 1 represents measurement occasion, level 2 represents subject, and level 3 represents center. Our approach is general with respect to sampling proportions, number of groups, and attrition rates over time. In addition, we also develop a cost model, as an aid in selecting the most parsimonious of several possible competing models. We derive sample size requirements (i.e., power characteristics) for a test of treatment-by-time interaction(s) for designs based on either subject-level or cluster-level randomization.

A Response Adaptive Repeated Measurement Design for Clinical Trials

❖ Kimmie C. Carriere, University of Alberta, 632 CAB, Dept Mathematical and Statistical Sciences, Edmonton, AB T6G 2G1 Canada, *kccarrie@ualberta.ca*; Yuanyuan Liang, University of Alberta

Key Words: response-adaptive design, multiple-objective designs, self and simple carryover, evaluation function, optimality criteria

In a response-adaptive design (RAD), we review and update the trial on the basis of outcomes in order to achieve a specific goal. Optimal designs for clinical trials are usually constructed under a single objective. In this paper, we develop a new adaptive allocation rule to improve the current strategies of building response-adaptive designs in two directions: (1) construct optimal multiple-objective designs to increase both the estimation precision and the proportion of patients assigned to a better treatment, and (2) use a more general model, that accommodates the self and simple mixed carry-over effects and random subject effect. In simulations we demonstrate that the designs constructed under the new proposed allocation rule are more efficient than fixed optimal designs in terms of the mean squared error.

Optimal Designs for Linear Mixed Effects Models

❖ Martijn P.F. Berger, University of Maastricht, P.O. Box 616, Maastricht, International 6200 MD Netherlands, *Martijn.Berger@stat.unimaas.nl*

Key Words: longitudinal designs, linear mixed effects models, optimal design, D-optimality, maximin

Longitudinal designs have been used to study effects over time in various different fields. These designs are characterized by the number and allocation of time points at which measurements are taken and by the number of distinct cohorts of people that are sampled. The most suitable models to analyze these data are the linear mixed effects models, because they enable random variation among and within subjects and cohorts. Efficient estimation of the parameters in these models depends on the characteristics of the longitudinal design. In this presentation we will discuss some of the problems that are involved in finding optimal designs for these models. D-optimal designs for these longitudinal designs will be presented and we will show how much efficiency is lost if sub-optimal designs are used. Guidelines for efficient designing of longitudinal studies will also be offered.

311 Modeling Time-Varying Exposures in Epidemiology ● ✪

Section on Statistics in Epidemiology, Section on Health Policy Statistics, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 2:00 pm–3:50 pm

Semiparametric Bayesian Latent Trajectory Models

❖ Amy H. Herring, The University of North Carolina at Chapel Hill, CB 7420, Chapel Hill, NC 27599, *amy_herring@unc.edu*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Bayesian, latent variable, exposure assessment, disinfection by-product, joint model

Exposures of interest often may vary over relatively long periods of time. Latent trajectory models (LTMs) characterize longitudinal exposure data using a finite mixture of curves. We address uncertainty in the number of latent classes and in the form of the class-specific curves using a semiparametric Bayesian approach, which is generalized to allow joint nonparametric modeling with a multivariate response. The proposed approach allows the response distribution to be unknown and varying with trajectory class. An MCMC algorithm is developed for posterior computation. The methods are motivated by an epidemiologic study of water quality and pregnancy outcomes and are compared to analyses using simpler exposure summaries commonly used in epidemiologic studies.

Bayesian Semiparametric Clustering of Functional Predictors

❖Jamie Bigelow, Duke University, *jbigelow@duke.edu*

Key Words: Dirichlet process, Functional data, hormones, species sampling

Motivated by the problem of classifying hormone trajectories, we propose a flexible semiparametric Bayesian methodology for hierarchical functional data. By choosing a species sampling random probability measure for the distribution of coefficients in a spline model, a general class of nonparametric Bayesian methods for clustering of functional data is developed. Allowing the spline basis to be unknown, one faces the problem of posterior simulation over a high-dimensional space of semiparametric models. To address this problem, we propose a novel Metropolis-Hastings algorithm for moving between models, with a nested generalized collapsed Gibbs sampler for updating the model parameters. Focusing on Dirichlet process priors for the distribution of the basis coefficients in multivariate linear spline models, we apply the approach to the problem of clustering of hormone trajectories.

Modeling Time-Varying Exposures in Epidemiology

❖Francesca Dominici, Johns Hopkins University, 615 N Wolfe St, Baltimore, MD 21205-2103, *fdominic@jhsph.edu*

Key Words: air pollution exposure, causal parameters, risk assessment

Despite the increasingly stringent national air quality regulations, adverse health effects associated with exposure to air pollution persists. Not surprisingly, regulators are looking for evidence of the gains in public health that have followed the implementation of costly regulatory policies. In this paper we develop statistical methods for assessing the health impact of air quality regulations. Specifically we plan to develop: 1) methods for predicting local and national PM long-term trends; 2) methods for cohort studies for estimating associations between long-term exposure to PM and mortality; and 3) a causal framework for estimating the total reduction of deaths attributable to regulations. The methods will be applied to the National Medicare Cohort Study (1999-2006).

312 Scagnostics ● ✪

Section on Statistical Graphics

Tuesday, July 31, 2:00 pm–3:50 pm

Scagnostics in R

❖Hadley Wickham, Iowa State University, 2519 Chamberlain Street, Apt 314, Ames, IA 50014, *h.wickham@gmail.com*; Duncan Temple Lang, University of California, Davis

Key Words: scagnostics, R, graphics

In this talk we discuss the implementation of calculation in R using R's C interface to call C++ code converted from Java. We will discuss graphical display and interaction possible with R graphics.

Scagnostic-Driven Autovisualization

❖Graham Wills, SPSS Inc., 233 S. Wacker Dr., Chicago, IL 60540, *gwills@spss.com*

Key Words: scagnostics, visualizations, statistical graphics

A common situation faced by data professionals is that of receiving a large table of data, often containing many variables, and only a vague idea of what to do with it. For example, a large table of data containing information about insurance customers might arrive with a note to find out which customers are unprofitable. Even with a more focused query, it is useful to get a good look at the data initially before building a model. In this paper we will present a tool designed to give an initial look at data sets by using scagnostics to build an informal model of the relationships between variables. Based on the scagnostic analysis, a multipart visualization is built that is designed to show the main features of the data set. The Scagnostic AutoVis tool will be demonstrated, its design discussed and further enhancements suggested.

Scagnostics for Projection Pursuit

❖Heike Hofmann, Iowa State University, 122 Snedecor Hall, Ames, IA 50014, *hofmann@iastate.edu*; Dianne Cook, Iowa State University; Hadley Wickham, Iowa State University

Key Words: exploratory data analysis, visualization, multivariate data

Finding "interesting" projections in high dimensional space has a long tradition, yielding in methodological solutions such as the grand tour (Asimov 1958) or projection pursuit methods. While the grand tour walks through the high-dimensional space on a path that covers all possible lower dimensional projections, this path is optimized in projection pursuit methods for one specific optimality index. We propose an application of graph-theoretic indices as optimization criterion in projection pursuit: instead of a single criterion, a combination of these indices allows us to look for projections that are e.g., 80% "skinny", 15% "clumpy" with 0% outliers, etc. By allowing to change the index setting interactively, the analyst can guide the projection pursuit in more ways than the usual parameters allow.

313 Bayesian Bioinformatics ●

Section on Bayesian Statistical Science, ENAR, Biometrics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Simultaneous Inference for Multiple Testing and Clustering via a Dirichlet Process Mixture Model

❖David B. Dahl, Texas A&M University, Department of Statistics, 3143, College Station, TX 77843-3143, *dahl@stat.tamu.edu*; Qianxing Mo, Texas A&M University; Marina Vannucci, Texas A&M University

Key Words: Bayesian nonparametrics, Correlated hypothesis tests, Model-based clustering, Multiple comparisons

We propose a Bayesian nonparametric model which exploits clustering for increased sensitivity in multiple hypothesis testing. We build on Dahl and Newton (2005) who showed that this was feasible by modeling the dependence among objects through clustering and then estimating hypothesis-testing parameters averaged over clustering uncertainty. We propose several improvements. First, we separate the clustering of the regression coefficients from the accommodation of heteroscedasticity. Second, our model accommodates more general experimental designs, such as permitting covariates and not requiring independent sampling. Third, we provide a more satisfactory treatment of nuisance parameters and hyperparameters. Finally, we do not require the designation of a reference treatment. The proposed method is compared in a simulation study to ANOVA and the BEMMA method of Dahl and Newton (2005).

A Bayesian Branching Process Model for Loss of Cell Cycle Synchrony

❖Edwin Iversen, Duke University, Box 90251, Institute of Statistics and Decision Sciences, Durham, NC 27708-0251, *iversen@stat.duke.edu*; David Orlando, Duke University; Allister Bernard, Duke University; Alexander Hartemink, Duke University

Key Words: branching process, deconvolution, cell cycle, flow cytometry, time course data

We describe a model for cell population dynamics in synchrony/release experiments wherein cells are released from synchrony and a time series of assay measurements is made on the population. Because the population loses synchrony, measurements represent a convolution over cell cycle states that depends on time since release. Our model characterizes the fraction of cells at any given time in each state; this data can be used to deconvolve population-level experimental measurements, such as temporal expression profiles, to reflect their true patterns of cell-cycle dependence. The model can be fit to budding fractions, flow cytometry measurements, or both. It captures population dynamics related to cell division using a branching process formulation. We describe the model, its relationship to a cell cycle time line and illustrate its utility in a study of S. cerevisiae.

Better Stochastic Mapping and Effective Tests for Evolutionary Innovation

❖ Marc A. Suchard, University of California, Los Angeles, Human Genetics, 695 Charles E. Young Dr., Los Angeles, CA 90095-1788, *msuchard@ucla.edu*; Vladimir N. Minin, University of California, Los Angeles

Key Words: Stochastic processes, Phylogenetics, MCMC, Evolution

Stochastic mapping is an invaluable tool to study evolutionary innovation. Previously successful examples include identifying developmental pathway differences between human and chimps and measuring the forces of selection on HIV over the course of infection. Current stochastic mapping tools derive from highly inefficient simulation techniques and are often misleading. Coaching evolutionary innovation as a novel Bayesian Markov Arrival Process, we are able to provide analytic expressions for the mean and variance of the expected number of events in an evolutionary history and recursive solutions to determine their probability mass function. From these expressions, we formulate a test for evolutionary innovation. Unlike simulation-based methods, test quality does not suffer as the number of traits and their complexity increases.

Associating High-Dimensional Response and Covariate Data

❖ Mahlet G. Tadesse, University of Pennsylvania, Department of Biostatistics & Epidemiology, 918 Blockley Hall, Philadelphia, PA 19104-6021, *mtadesse@cceb.med.upenn.edu*; Stefano Monni, University of Pennsylvania

Key Words: Markov chain Monte Carlo, high-dimensional data

In recent years, high-dimensional datasets have become common in many applications. Several methods have been developed to exploit the wealth of information in these data. Most of the procedures have focused on relating the high-dimensional covariates to univariate outcomes, or at most a few correlated outcomes. Often, however, both the response and covariate data are high-dimensional. We propose a method to identify subsets of covariates associated with sets of correlated outcomes. We illustrate the method with an application to eQTL analysis, in which gene expression microarray data are related to genotype data from thousands of SNP markers. The method allows the identification of correlated gene expression patterns modulated by sets of DNA variations.

314 Issues in Modeling Product Choices in Marketing ● ✪

Section on Statistics and Marketing

Tuesday, July 31, 2:00 pm–3:50 pm

Market Share Constraints and the Loss Function in Choice-Based Conjoint Analysis

❖ Timothy Gilbride, University of Notre Dame, 315 Mendoza College of Business, Notre Dame, IN 46530, *tgilbrid@nd.edu*; Peter Lenk, The University of Michigan; Jeff Brazell, The Modellers, LLC

Key Words: Hierarchical Bayes, Loss Function, Bayesian Decision Theory, Conjoint Analysis

Choice-based conjoint analysis is a popular marketing research technique to learn about consumers' preferences and to make market share forecasts. Managers expect these forecasts to be "realistic" in terms of being able to replicate market shares at some pre-specified or "base case" scenario. Frequently, there is a discrepancy between the forecasted and base case market share. This paper presents a Bayesian decision theoretic approach to incorporating base case market shares into conjoint analysis via the loss function. We specify an appropriate loss function and all estimates are formally derived via minimizing the posterior expected loss. We contrast this approach to using informative priors. MCMC methods for performing the analysis are presented. The approach is demonstrated with simulated data and actual market research studies using both multinomial logit and probit models.

Unobserved Change in Conjoint Analysis

❖ Thomas Otter, The Ohio State University, Department of Marketing & Logistics, 2100 Fisher Hall, Columbus, OH, *otter.2@osu.edu*; Sylvia Fruehwirth-Schnatter, Johannes Kepler Universitaet Linz; Regina Tuechler, Vienna University of Economics and Business Administration

Key Words: conjoint analysis, change point, heterogeneity, Bayesian Analysis

We develop models of change that are parsimonious in the sense that each respondent is characterized by one vector of regression weights, like in the standard mixed effects model. In the first model every respondent's observation error variance is allowed to change according to a change point process. The change points are modeled as random effects. In the second model change is viewed to affect which attributes are being used by a respondent in forming product evaluations. This may be expressed as a dynamic variable selection problem where indicators pointing to the active variables evolve over the course of the experiment. We model these indicators as random effects that evolve according to a change point process. We conclude with a comparison to an approach where observations separated by change points are independent (conditional on hyper-parameters).

Deconstructing Each Item's Category Contribution

❖ Sharad Borle, Rice University, Jones Graduate School of Management, 6100 Main Street, Houston, TX 77005, *sborle@rice.edu*; Peter Boatwright, Carnegie Mellon University; Kirthi Kalyanam, Santa Clara University

Key Words: Apparel Retailing, Retail assortment, Category Management, out-of-stocks, hierarchical Bayes, COM-Poisson

This paper provides an empirical study of the role of every item in an assortment. Our results show many items affect category sales over and above their own sales volume. After deconstructing the role of a stockout of individual items into three effects, lost own sales, substitution to other items, and the category sales impact, we find that the category impact has the largest magnitude. Interestingly almost every single individual item affects category sales (it is not restricted to top selling items); variety indeed is the price of entry in retailing. Our results support recent findings that more frequently purchased categories are less adversely affected by reductions in assortment. We also find that the assortment appears to gain attractiveness when certain items are out of stock, a result that is consistent with the discussion in the literature concerning category clutter.

315 Investing in Our Youth as Future Users of Statistics and Future Respondents

Section on Government Statistics, Section on Statistical Education

Tuesday, July 31, 2:00 pm–3:50 pm

Statistics Canada's Education Outreach Program: Fostering a Data and Knowledge Culture in Canadian Youth

❖Mary Townsend, Statistics Canada, 13 C7 Jean Talon Building, 175 Tunneys Pasture Parkway, Ottawa, ON K1A 0T6 Canada, *mary.townsend@statcan.ca*

Key Words: statistical literacy, data management skills development, learning statistics

Through Education Outreach, Statistics Canada makes its information relevant and understandable to young Canadians, so that they become comfortable using it to make informed decisions today as students, and in the future as parents, business people and respondents. We accomplish this both online, through our Learning Resources website that offers relevant data and expertise, and at a grassroots level, through a network of teacher champions who provide data management workshops to teachers and students. We also encourage our own employees to bring their considerable knowledge to local schools. This paper highlights the award-winning initiatives of Statistics Canada's Education Outreach Program that engage youth in active statistical learning. By investing in Education Outreach, Statistics Canada is paving the way to establish a data and knowledge culture in Canada.

How to Make the Census Exciting for Educators and K–12 Students

❖Kimberly Crews, U.S. Census Bureau, Public Information Office, 4600 Silver Hill Road, Room 8H052, Washington, DC 20233-0900, *kcrews@census.gov*

Key Words: Census Bureau, Education, Children, data user, United States

The U.S. Census Bureau produces a wealth of information about the people and economy of the United States. For the Census Bureau to continue to produce high quality statistical information, it must maintain a high level of cooperation among census and survey participants. Nationwide, there are some 55-million school-age children in the United States who one day will be asked to participate in the census or other survey. They are also current and future data users. This session highlights the activities of the Census in Schools program that facilitates the use of census information by providing engaging teaching materials and training to educators across the curriculum in support of the 2010 Census.

Web Services as an Aid in Promoting Statistical Knowledge Within Educational Institutions

❖Reija Helenius, Statistics Finland, P.O.Box 3A, FI-00022 Statistics Finland, Helsinki, Finland, *reija.helenius@stat.fi*

Key Words: Finland, Statistics Finland, Educational institution, Statistical literacy, Web services, Web based learning material

Promotion of skills in statistical literacy and in the use of statistics within educational institutions is one of the strategic goals for Statistics Finland, similarly as for other statistical offices. Statistical literacy requires practice and knowledge of the basic facts and concepts related to statistics. Statistics Finland's co-operation with teachers has shown that educational institutions have a distinct need for statistical information and advice on teaching how to use it. Statistics Finland stresses the importance of developing web services for educational institutions. Teachers have actively participated in Statistics Finland¥s web product development both as members of the product development teams and as commentators on the products.

316 Causal Inference in Education Research ● ✪

Social Statistics Section, Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Tuesday, July 31, 2:00 pm–3:50 pm

Random Assignment Designs, Issues of Implementation, and Analytic Challenges Faced in Education Settings

❖David Myers, American Institutes for Research, 1000 Thomas Jefferson St. NW, Washington, DC 20007, *dmyers@air.org*

Key Words: causal inference, clinical trial, randomized experiment

In recent years there has been a push towards randomized experiments as the best way to determine "what works" in education. This talk will present examples of randomized experiments in education, including a cluster randomized trial of remedial reading programs, a nationally representative trial of the Upward Bound program, and a charter schools evaluation. When done well, randomized experiments can yield unbiased estimates of program effectiveness. However, in practice there are a number of challenges to doing random assignment studies in education, including recruitment and study retention, noncompliance, and clustering of students within classrooms, schools, and neighborhoods. Many of these issues lead to methodological challenges that have not yet been fully solved.

Causal Inference and Accountability in NYC Public Schools

❖Jennifer Hill, Columbia University, New York, NY 10027, *jh1030@columbia.edu*

Key Words: accountability, causal inference, schools

Accountability is the latest battle cry in the war on the failure of our educational system in the urban US. This fight was taken up and turned into national policy by No Child Left Behind. But what does accountability mean when viewed from a causal perspective? And how can principles regarding causal inference influence the way that accountability policies are manifested in large US city school districts? This question is pursued and one solution described in the context of the New York City Public School System.

Regression Discontinuity Analyses of the Effects of NCLB Accountability Provisions on Student Achievement

❖John R. Lockwood, RAND Corporation, 4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213, *lockwood@rand.org*; Brian P. Gill, RAND Corporation; Messan C. Setodji, RAND Corporation; Francisco Martorell, RAND Corporation

Key Words: No Child Left Behind, longitudinal achievement data, school accountability, quasi-experimental analysis

The accountability provisions of the federal No Child Left Behind (NCLB) Act stipulate that each year, schools must meet academic proficiency targets in reading and mathematics or face sanctions and interventions aimed at improving student achievement. The academic proficiency targets determining whether schools make "adequate yearly progress" involve the percentages of all students, and of certain subgroups of students, who reach proficiency on annual state accountability assessments. These thresholds imply that schools are assigned to NCLB "treatment" groups on the basis of a cutscore on a continuous variable, providing the opportunity for regression discontinuity analyses of the impacts of the interventions. In this talk we discuss the application of these methods to both school-level and student-level measures of achievement to assess the impacts of certain NCLB provisions.

317 Statistical Methods in Nanotechnology ● ✪

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Tuesday, July 31, 2:00 pm–3:50 pm

Experimental Design and Analysis for Robust Synthesis of Nanostructures

❖ Chien-Fu Jeff Wu, Georgia Institute of Technology, School of Industrial and System Engineering, 765 Ferst Drive, Atlanta, GA 30332-0205, *jeffwu@isye.gatech.edu*; Tirthankar Dasgupta, Georgia Institute of Technology; Roshan Vengazhiyil, Georgia Institute of Technology; Zhong Lin Wang, Georgia Institute of Technology; Christopher Ma, Georgia Institute of Technology

Key Words: design of experiments, nano-materials, robust parameter design, GLM

An effort is made to systematically investigate the best process conditions that ensures synthesis of different types of one dimensional cadmium selenide nanostructures with high yield and reproducibility. Through a designed experiment and rigorous statistical analysis of experimental data, models linking the probabilities of obtaining specific morphologies to the process variables are developed. A new iterative algorithm for fitting a Multinomial GLM is proposed and used. The optimum process conditions, which maximize the above probabilities and make the synthesis process less sensitive to variations of process variables around set values are derived from the fitted models using Monte-Carlo simulations. Some future research directions on experimental design and analysis of experimental data are highlighted.

Challenges Related to Reliability and Statistics in Nano Electronics

Way Kuo, University of Tennessee; ❖ Tao Yuan, University of Tennessee, Department of Industrial and Information Engineering, Knoxville, TN 37996-0700, *tyuan@utk.edu*

Key Words: reliability, nano electronics

To Build for the Future, we must achieve major advances related to reliability in addition to exploring and discovering interdisciplinary connections in important cutting-edge research areas. The technologies for today's design and manufacturing have for some time been steadily moving from the realm of the micro- to the nano scale, but advancements in reliability and statistics have not kept up with the pace. In this presentation, we will talk about modeling and analysis for reliability of nano electronics.

318 Anomaly Detection Methods for Multivariate, Distributed, and Stratified Data in Biosurveillance ● ✪

Section on Statisticians in Defense and National Security, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Section on Risk Analysis, Section on Physical and Engineering Sciences

Tuesday, July 31, 2:00 pm–3:50 pm

A Multivariate Bayesian Method for Spatial Biosurveillance

❖ Daniel B. Neill, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, *neill@cs.cmu.edu*; Andrew W. Moore, Carnegie Mellon University; Gregory F. Cooper, University of Pittsburgh

Key Words: cluster detection, data mining, biosurveillance

We present a new method for spatial cluster detection, the "multivariate Bayesian scan statistic" (MBSS). MBSS combines information from multiple data streams in a Bayesian framework, computing the posterior probability of each outbreak type in each spatial region. MBSS learns a Gamma-Poisson model for each region from historical data, and models the effects of each outbreak type on each stream using expert knowledge. We evaluated MBSS on simulated outbreaks injected into three streams of OTC sales data (cough, fever, thermometers) for western PA. At 1 fp/month, MBSS detected outbreaks in 1.59 days, as compared to 2.23, 2.59, and 1.99 days for univariate Bayesian detectors monitoring each stream. MBSS was also able to characterize outbreaks by pinpointing the affected region, and accurately distinguished between different outbreak types. This work was supported by NSF grant IIS-0325581.

Multivariate Health Surveillance Using a Network of Bayesian Data Fusion Models

❖ Zaruhi R. Mnatsakanyan, Johns Hopkins University, 11100 Johns Hopkins Road, Applied Physics Laboratory, Laurel, MD 20723, *zaruhi.mnatsakanyan@jhuapl.edu*; Howard S. Burkom, Johns Hopkins University

Key Words: Bayesian Belief Networks, Data Fusion, Decision Support System, Anomaly Detection, False Positive Reduction, Biosurveillance algorithms

The model is a network of distributed Bayesian Network (BN) for region-wide surveillance of respiratory outbreaks using multiple data sources. Each BN combines inputs of alerting algorithms and preprocessed data to estimate the probability of true health anomalies at reduced alert rates. Each also discriminates between seasonal influenza and other significant events. The BN structures are based on association of anomalies with different data types and correlation of those events based on the decision-making logic of domain experts. The model detected Maryland influenza outbreaks in 2003-04, 2004-05 and did not alert for the 2005-06 season, which saw no significant rise in flu. The network fusion yields epidemiological false alarm rates significantly below those of univariate temporal alerting algorithms of ESSENCE surveillance systems.

Epidemiological Network Models for Disease Outbreak Detection

❖Ben Y. Reis, Harvard Medical School, 1 Autumn St, Room 540.1, Boston, MA 02115, *reis@mit.edu*

Key Words: networks, surveillance, detection, pandemics, public events

We describe a class of epidemiological network models that rely on monitoring the relationships among different healthcare data streams, instead of monitoring the data streams themselves. The relational nature of these models has the potential to increase their robustness to unpredictable baseline shifts. Simulation results show that the network models provide better detection of localized disease outbreaks, and greater robustness to healthcare utilization shifts than a standard modeling approach. This integrated network model of epidemiological data streams and their inter-relationships has the potential to improve current public health surveillance efforts, providing more reliable localized outbreak detection and situational awareness in the face of shifts in healthcare utilization during epidemics and major public events.

KFC: Anomaly Detection in Multivariate Cross-Classified Datastreams

❖Colin Goodall, AT&T Labs - Research, 200 S Laurel Ave, D4 3D28, Middletown, NJ 07748, *cgoodall@att.com*

Key Words: Multivariate anomaly detection, Kalman Filter, Ratio shrinkage estimation, Loglinear and graphical models, Biosurveillance, Bayesian models

Consider daily counts of emergency department visits, cross-classified into a list of counts by medical syndrome, hospital, patient age group, zip, and gender. We monitor these many time series for anomalies, such as spikes in visits related to illnesses in specific sub-populations. The Kalman Filter for Contingency Tables, or KFC, provides a framework and specific methodology for modeling and anomaly detection. It includes a loglinear or graphical model for expected counts, which are updated by a Kalman Filter, and shrinkage estimation applied to ratios of observed to expected, adapted from DuMouchel's Gamma Poisson Shrinker for very large tables of counts. KFC is part of ongoing work at AT&T Labs - Research in collaboration with Emergency Medical Associates of New Jersey, and extends new methods for biosurveillance of univariate time series to highly multivariate settings.

319 Integrating Innovative Technology Resources into Introductory Courses

Section on Statistical Education
Tuesday, July 31, 2:00 pm–3:50 pm

Integrating Innovative Technology Resources into Introductory Courses

❖Jacqueline Miller, The Ohio State University, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *jbm@stat.ohio-state.edu*; ❖Kim Robinson Gilbert, Georgia Gwinnett College, 1000 University Center Lane, Lawrenceville, GA 30043, *kimrobinson@clayton.edu*; ❖Nicholas Horton, Smith College, Department of Mathematics and Statistics, Clark Science Center, Northampton, MA 01063-0001, *nhorton@email.smith.edu*; ❖Ginger Holmes Rowell, Middle Tennessee State University, Department of Mathematical Sciences, PO Box 34, Murfreesboro, TN 37130, *rowell@mtsu.edu*; ❖Roxy Peck, California Polytechnic State University, Building 25, Room 107D, San Luis Obispo, CA 93407, *rpeck@calpoly.edu*

Key Words: Education, Technology, Introductory course, Assessment

Technology has had a profound effect on instruction in statistics courses. Although many instructors are comfortable using graphing calculators and computer software, other underutilized resources are available. Each panelist will describe how they integrate some of the more specialized technology resources, such as clickers, PC Tablet, online ARTIST assessments, CAUSEweb, and an online survey. Discussion and questions will be encouraged.

320 Real-Time Analysis of Revised Macroeconomic Data ●

Business and Economics Statistics Section
Tuesday, July 31, 2:00 pm–3:50 pm

Real-Time Filtered Estimates of Final Superlative CPI Based on an Estimated Monthly VAR Model of Initial and Final Superlative CPI

❖Peter Zadrozny, Bureau of Labor Statistics, 2 Massachusetts Ave, NE Room 3105, Washington, DC 20212, *zadrozny.peter@bls.gov*

Key Words: Kalman filtering, missing data, revised data, reduced-form modelling

Initial SCPI estimates are released every month with 1 month delay. Final SCPI estimates are released every February for all months 2 years ago, with 15-25 months delay. Estimates are in historical form (bivariate, compact, with no missing values) when indexed by months of occurrence and are in real-time form (50-variate, sparse, with many missing values) when indexed by months of release. A bivariate VAR model is estimated using historical-form data. A Kalman filter, which can handle missing data, is applied to the estimated model in state-space form and the released estimates in real-time form in order to compute filtered estimates of final SCPI. Applied to recent U.S. data in log form, the method produces RMSE < 1% of filtered estimates of final SCPI, compared with RMSE < 20% of initial releases of SCPI, as estimates of final releases of SCPI.

Real-Time Estimation of Revised GDP Based on an Estimated Model of Initial and Revised GDP

❖Baoline Chen, Bureau of Economic Analysis, 1441 L Street NW, Washington, DC 20230, *baoline.chen@bea.gov*; Peter Zadrozny, Bureau of Labor Statistics

Key Words: Kalman filtering, Missing data, revised data, reduced form modeling

This paper develops and illustrates with U.S. GDP a two-step method based on an estimated VAR model for estimating in any period the final GDP for that period. Step one estimates the model for included variables by applying either OLS or maximum likelihood using three initial releases and three annual revisions of GDP from 1977 to 2003, indexed either by periods to which they pertain or by periods in which they are released. Step two estimates final GDP in every period using only sample information which is available through that period. This can be done by applying the missing-data Kalman filter to the estimated model using data in sparse real-time form. In the application, filtered estimates of final GDP are compared to eventually released final values in terms of RMSE measures of closeness.

VAR Estimation and Forecasting When Data Are Subject to Revision

❖Evan Koenig, Federal Reserve Bank of Dallas, Research Department, 2200 North Pearl Street, Dallas, TX 75201, *evan.f.koenig@dal.frb.org*; N. Kundan Kishor, Texas Tech University

Key Words: data revisions, forecasting, Kalman filter, vector autoregression

Conventional VAR estimation and forecasting ignores the fact that economic data are often subject to revision many months or years after their initial release. This paper shows how VAR analysis can be modified to account for such revisions. The proposed approach assumes that government statistical releases are efficient with a finite lag. It takes no stand on whether earlier revisions are the result of "news" or of a reduction in "noise." The technique is illustrated using several forecasting models of real economic activity. In each case, the proposed procedure outperforms conventional VAR analysis and the more-restrictive methods for handling the data-revision problem that are found in the existing literature. It yields forecasts of GDP growth and the unemployment rate that are more accurate than those from the Blue Chip newsletter and the Survey of Professional Forecasters.

Modeling Data Revisions: Measurement Error and Dynamics of 'True' Values

❖Jan Jacobs, University of Groningen, PO Box 800, Groningen, 9700 AV Netherlands, *j.p.a.m.jacobs@rug.nl*; Simon van Norden, HEC Montreal

Key Words: real-time data analysis, data revisions

Realistic modeling of data revisions can play an important role in policy formulation. A common way to model data revisions is to set up a state-space model with separate blocks for measurement errors and the dynamics of "true" values. However, empirical work suggests that measurement errors typically have much more complex dynamics than such models allow. This paper describes a state-space model with richer dynamics in these measurement errors, including the noise, news and spillover effects documented in this literature. The result is a unified and flexible framework that allows for more realism in the model of data revision and also allows analysis of optimal real-time estimation of trends and cycles in real time. We illustrate the application of this framework with an analysis of a real-time data set on U.S. real output.

A State Space Approach to the Policymaker's Data Uncertainty Problem

❖Jana Eklund, Bank of England, Threadneedle Street, London, EC2R 8AH, United Kingdom, *jana.eklund@bankofengland.co.uk*; Alastair Cunningham, Bank of England; George Kapetanios, Bank of England/ Queen Mary, University of London; Chris Jeffery, Bank of England; Vincent Labhard, European Central Bank

Key Words: Real-time data analysis, State space models, Data uncertainty, Data revisions

The paper describes the challenges that uncertainty over the true value of key macroeconomic variables poses for the policymaker and the way in which she may form and update her view of the evolution of economic time-series in light of a range of indicators and models. Specifically, it casts the data uncertainty challenge in state space form and describes a two-step estimator for the resulting signal extraction problem. Real-time data are first used to estimate the statistical properties of past revisions to published estimates of macroeconomic variables; and these properties are then imposed in maximum likelihood estimation of the full state space model.

321 The Role of Statistics in Ecological and Climate Modeling ● ✪

Section on Statistics and the Environment, ENAR, International Indian Statistical Assoc, Reps. For Young Statisticians

Tuesday, July 31, 2:00 pm–3:50 pm

Getting What We Need from Wireless Sensor Networks: A Role for Inferential Ecosystem Models

❖James Clark, Duke University, Nicholas School of the Environment, Durham, NC 27708, *jimclark@duke.edu*; Pankaj Agarwal, Duke University; David Bell, Duke University; Carla Ellis, Duke University; Paul Flikkema, Northern Arizona University; Alan E. Gelfand, Duke University; Gabriel Katul, Duke University; Kamesh Munagala, Duke University; Gavino Puggioni, Duke University; Adam Silberstein, Duke University; Jun Yang, Duke University

Key Words: environmental change, prediction, hierarchical Bayes, wireless networks, forests

Efficient wireless sensor networks require inferential ecosystem models that can weigh the value of an observation against cost of transmission. Transmission costs make observations 'expensive'; networks are deployed in remote locations without access to power. The capacity to sample intensively makes sensor networks powerful, but high frequency data have value only at specific times and locations. Given that intensive sampling is sometimes critical, but more often wasteful, how do we control transmission? The value of an observation can be evaluated in terms of its contribution to estimates of state variables and parameters. Network control must be dynamic and driven by models capable of learning about both the environment and the network. Inference is needed to weigh the contributions against transmission cost. We discuss Bayesian inference to control network transmission.

The Uncertain Hockey Stick: A Statistical Perspective on the Reconstruction of Past Temperatures

❖Douglas Nychka, National Center for Atmospheric Research, Box 3000, Boulder, CO 80307-3000, *nychka@ucar.edu*

Key Words: Hockey Stick, uncertainty, ensembles, NH temperature

A reconstruction of past temperatures based on proxies is inherently a statistical process and a deliberate statistical model for the reconstruction can also provide companion measures of uncertainty. we suggest a framework where the reconstruction is expressed as a conditional distribution of the temperatures given the proxies. Random draws from this distribution provide an ensemble of reconstructions where the spread among ensemble members is a valid statistical measure of uncertainty. This approach is illustrated for Northern Hemisphere temperatures and the multi-proxy data used by Mann, Bradley and Hughes (1999). Here we explore the scope of the statistical assumptions needed to carry through a rigorous analysis and use Monte Carlo sampling to determine the uncertainty in maxima or other complicated statistics in the reconstructed series.

Modeling Extremes in Regional Climate Model Simulations

❖Daniel Cooley, Colorado State University, 321 Judson St, Longmont, CO 80501, *cooleyd@stat.colostate.edu*; Stephan Sain, National Center for Atmospheric Research

Key Words: precipitation, spatial statistics, Bayesian hierarchical model

Regional climate models (RCM) are designed to capture local climate behavior that general circulation models (GCM) cannot because of their course resolution. Like GCMs, RCMs are tuned so that the weather they simulate resembles that which is currently observed under current climate conditions. Although the RCM-simulated weather pretty closely approximates local climate, there is some question as to whether they accurately represent extreme weather. In this work, we build a Bayesian hierarchical model for the extreme precipitation data from a RCM for the Western United States. We compare the extremes of a control run of the RCM to those of a run simulated under increased carbon dioxide scenarios. We also compare the extremes produced by the RCM to those recorded at weather stations over the last half of the twentieth century.

Functional ANOVA Modeling of Regional Climate Model Experiments

❖ Cari Kaufman, Statistical and Applied Mathematical Sciences Institute/North Carolina State University, 19 TW Alexander Drive, PO Box 14006, Research Triangle Park, NC 27709, *cgk@samsi.info*; Stephan Sain, National Center for Atmospheric Research

Key Words: climate models, functional data, spatial data

The North American Regional Climate Change Assessment Program (NARCCAP) is exploring uncertainties in regional climate modeling over North America by nesting a set of regional models within a set of global models. As a diagnostic tool to compare models across spatial regions, we propose a functional ANOVA approach, in which we decompose the climate response into a common mean response, effects due to regional model, effects due to global model, and interactions. We model these functional effects using Gaussian process priors, and we use the posterior distribution to compare sources of variability in model output.

Spatio-Temporal Processing of MISR's Aerosol Optical Depth Data

❖ Tao Shi, The Ohio State University, 317 Cockins Hall, 1958 Neil Ave., Department of Statistics, Columbus, OH 43210, *taoshi@stat.ohio-state.edu*; Noel Cressie, The Ohio State University

Key Words: data processing, Kalman filtering, kriging, spatial heterogeneity

NASA's Multiangle Imaging Spectro Radiometer (MISR) provides level-2 data at 1.1km by 1.1km spatial resolution, and the satellite returns to its original location on the globe every 16 days. These level-2 data can be converted to level-3 data at a much lower spatial resolution by averaging those observations falling in lower-resolution pixels over a contiguous number of days; this results in a level-3 monthly data product at a resolution of 0.5 degree latitude by 0.5 degree longitude. Our goal is to fill in missing data and to de-noise the existing data, at level-3, in a statistically optimal way, where we exploit both the spatial and temporal dependencies in the data. We propose a method that combines elements of kriging and Kalman filtering to accomplish these goals, and we shall use the MISR Aerosol Optical Depth data to illustrate our approach.

322 Advanced Quantitative Brain MRI ● ✪

Biometrics Section

Tuesday, July 31, 2:00 pm–3:50 pm

The Acquisition and Statistical Analysis of High Temporal Resolution 3D fMRI Data

❖ Martin Lindquist, Columbia University, 1255 Amsterdam Ave, MC 4690, New York, NY 10027, *martin@stat.columbia.edu*

Key Words: fMRI, brain, imaging, high frequency

In this work we present a new acquisition and analysis technique for performing rapid 3D fMRI. Our acquisition strategy is based on repeatedly sampling a small central region of 3D k-space every 100ms, allowing us to obtain a low spatial resolution image with extremely high temporal resolution (compared to the standard 2000ms). The increased resolution allows one to effectively study the intricacies of the hemodynamic response function (HRF) which allows for accurate estimates of differences in activation timing across the brain. The feasibility and efficiency of the approach is confirmed using data from a visual-motor task. The results indicate that important timing information can be found by studying the initial negative dip. As the dip occurs in a time scale closer to that of the neural activity it may provide more accurate estimates of timing compared to more standard approaches.

Semiparametric Detection of Significant Activation for Brain fMRI

❖ Chunming Zhang, University of Wisconsin-Madison, Department of Statistics, 1300 University Avenue, Madison, WI 53706, *cmzhang@stat.wisc.edu*; Tao Yu, University of Wisconsin-Madison

Key Words: Deconvolution, Local polynomial regression, Nonparametric test, Spatial temporal data, Stimuli, Time resolution

Functional MRI (fMRI) aims to locate activated regions in human brains when specific tasks are performed. The conventional tool for analyzing fMRI data applies some variant of the linear model, which is restrictive in modeling assumptions. To yield more accurate prediction of the time-course behavior of neuronal responses, the semiparametric inference for the underlying hemodynamic response function is developed to identification of significantly activated voxels. We demonstrate that a class of the proposed semiparametric test statistics, based on the local-linear estimation technique, follow chi-squared distributions under null hypotheses for a number of useful hypotheses. Simulation evaluations and real fMRI data application endorse that the semiparametric inference procedure delivers more efficient detection of activated brain areas than popular imaging analysis tools.

Brain Asymmetry Analysis via Weighted Spherical Harmonic Representation

❖Moo K. Chung, University of Wisconsin-Madison, Department of Statistics, 1300 University Ave, Madison, WI 53708, *mchung@stat.wisc.edu*

Key Words: Brain, Image Analysis, Spherical Harmonics, Symmetry, Asymmetry, SPHARM

Previous approaches on brain asymmetry analysis start with mirroring 3D magnetic resonance images. The correspondence across the hemispheres is established by registering the original image to the mirrored image. We present a radically different approach that utilizes the recently developed weighted spherical harmonic (SPHARM) representation. The weighted-SPHARM is a partial differential equation (PDE) based smoothing technique where the solution of the PDE is expressed as a weighted linear combination of harmonics. Due to the angular symmetry in the spherical harmonics, the correspondence across can be obtained without the time consuming image registration. Our approach can reduce the Gibbs phenomenon by making the coefficients of the expansion exponentially decaying faster and thus making more rapid convergence for discontinuous measurements.

Estimation of Pre-Post Surgical Changes in the fMRI Visual Field Map

❖ Raymond Hoffmann, Medical College of Wisconsin, 1210 Larchmont Drive, Waukesha, WI 53186, *hoffmann@mcw.edu*; Edgar A. DeYoe, The Medical College of Wisconsin; Mary Jo Maciejewski, Marquette University; Daniel Rowe, Medical College of Wisconsin

Key Words: fMRI, Spatial Statistics, Bayesian, image analysis

The visual field map is produced by mapping active voxels of the visual cortex onto a circular disk corresponding to the points of a circular visual target. The location of areas on the target can be identified by the temporal course of the fMRI (functional Magnetic Resonance Image) signal in the visual cortex. Each scan (pre-surgical and post-surgical) produces a different set of points due to due to noise and variability in the positioning of the head between scans. The observed data is a non-homogeneous 2-dimensional point process with higher density in the center corresponding to the larger area of the visual cortex that is connected to the center of the retina. The underlying process intensity is estimated both by a 2D nonparametric smoother and a non-parametric 2D Dirichlet prior on the visual field. The posterior density is used to estimate the probability of change.

Statistical Analysis of Diffusion Tensors in Diffusion-Weighted Magnetic Resonance Image Data

❖Hongtu Zhu, The University of North Carolina at Chapel Hill, Department of Biostatistics, McGavran Greenberg Hall CB7420, Chapel Hill, NC 27599-3804, *hzhu@bios.unc.edu*; Heping Zhang, Yale University; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill; Bradley G. Peterson, Columbia University

Key Words: Diffusion tensors, Eigenvalue, Eigenvector, Pseudo-likelihood rato, Weighted least square

Diffusion tensor imaging has been widely used to construct the structure and orientation of fibers in biological tissues. The aim of this study is to provide a comprehensive theoretical framework of statistical inference for quantifying the effects of noise on diffusion tensors, on their eigenvalues and eigenvectors, and on their morphological classification. We propose a semiparametric model parametric model to account for noise in diffusion-weighted images. We then develop a one-step, weighted least-squares estimate of the tensors and justify use of the one-step estimates based on our theoretical framework and computational results. We also quantify the effects of noise on the eigenvalues and eigenvectors of the estimated tensors by establishing their limiting distributions. We construct pseudo-likelihood ratio statistics to classify tensor morphologies.

323 Challenges in Medical Device Trials: Novel Approaches to Design and Analysis ● ✪

Biopharmaceutical Section, Biometrics Section, ENAR

Tuesday, July 31, 2:00 pm–3:50 pm

Experimental Design for Assessing the Effect of Bone Density on Mechanical Properties After Vertebroplasty

❖Chul Ahn, Food and Drug Administration, 1350 Piccard Dr, Rockville, MD 10850, *chul.ahn@fda.hhs.gov*

Key Words: Spinal surgery, Osteoporosis, Graeco-Latin Square, Mixed model, Experimental Design

This is an experimental design problem regarding spinal surgery to repair fractured vertebral bodies with acrylic cement. The question is how to assign the samples to the various treatment groups and choose the correct statistical model to analyze. The experimenter is injecting cadaver spinal bones with acrylic cement to see how bone density and cement dose affect the mechanical properties. There are 5 treatment groups (positive control, negative control, 5%, 15% and 30% cement fill), and 13 cadavers, each of which has a unique bone mineral density that the experimenter would like to correlate with the response variables. Each cadaver yields 5 thoracic and 5 lumbar vertebrae, and the experimenter wishes to randomly distribute these 10 vertebrae into the 5 treatment groups, since the specific level of each vertebra may affect the mechanical properties.

On the Use of Fisher's Exact Test When Randomization Is Sitewise

❖Shanti Gomatam, Food and Drug Administration, 1350 Piccard Dr HFZ 550, CDRH, Division of Biostatistics CDRH, Rockville, MD 20850, *shanti.gomatam@fda.hhs.gov*

Key Words: center-specific randomization, clinical trials, Exact test

Fisher's exact test is often used in the analysis of data from clinical trials. However, whereas Fisher's exact test assumes overall randomization, randomization implementation in most trials is center-specific. Does this make a difference? This question is addressed here.

Control of Type I Error and the Correlation of Multiple Endpoints in a Medical Device Trial

❖Kevin Najarian, Boston Scientific Corporation, 41 Howard Street, Milford, MA 01757, *najariak@bsci.com*; Carola Alfaro, Boston Scientific Corporation

Key Words: Medical Device, Multiple Endpoints, Type I Error Rate, Multiplicity, Correlation, Homogeneity

Medical device clinical trials often consider multiple endpoints to assess the efficacy and safety of a given treatment. And while these clinical response endpoints may be classified as primary and secondary, the interpretation of the results must be considered carefully. Multiple univariate analyses can inflate the type I error rate, and therefore, some adjustments of the observed p-values for multiplicity of tests may be necessary. In addition to many common techniques for multiplicity adjustments, there are some that utilize correlation information and consider the homogeneity of treatment effects to determine the magnitude of the adjustments necessary. An analysis of real-life simulated medical device data is considered to assess techniques for multiplicity with added regard to the relationships among endpoints.

A Comparison of Statistical Approaches for Long-Term MACE Rate in a Medical Device Trial

❖Jian Huang, Boston Scientific Corporation, One Scimed Place, Maple Grove, MN 55311, *jhuangmd@yahoo.com*; Lan Pan, Boston Scientific Corporation

Key Words: Bias, Binary Data, Informative Dropout, Relative Risk, Simulation

Binary data can be estimated the treatment effect in a variety of ways. There is a growing concern about binary rate estimation when analyzed without taking into consideration the dropout in the long term safety trials. In this presentation we compared the following seven statistical approaches for determining a appropriate method to handle dropout during a long term follow-up: Binary rate based on ITT, Binary rate based on complete case, Event rate adjusting denominator based on adequate follow-up, Binary rate with single imputation, Binary rate with opposite imputation, Survival analysis for time to first event, and Events per person years. The relative risk estimate, bias, and statistical power were evaluated via Simulation. Simulation results suggest that time to event survival analysis and events per person years method generally are more appropriate for a long term MACE rate.

The SYNTAX Score: Development of a Predictive Scoring Algorithm To Determine Treatment Assignment in Complex Coronary Patients

❖Eric Bass, Boston Scientific Corporation, 1400 Hamline Avenue, South Town Square MS 9-315, St. Paul, MN 55112, *basse@bsci.com*; Peggy Pereda, Boston Scientific Corporation

Key Words: Predictive Modeling, Logistic Regression

The primary goal of the current trial is to assess the optimal revascularization treatment for patients with de novo three-vessel (3VD) or left main (LM) coronary artery disease by randomizing patients to either percutaneous coronary intervention (PCI) with drug eluting stent (DES) or coronary artery bypass grafting (CABG). An additional objective of the trial is to develop a SYNTAX Score by defining a combination of clinical risk factors and angiographic/anatomical characteristics that most closely predict the patient's probability of having a Major Adverse Cardiac and Cerebrovascular Event (MACCE) up to 1 year. From these parameters a risk score will be derived. Univariate logistic regression and proportional hazard modeling methods will be examined. Additionally, multivariate logistic regression and proportional hazard modeling will be conducted with multiple selection methods.

324 Statistical Issues in HIV/AIDS Research

ENAR, Biometrics Section, Biopharmaceutical Section, Section on Health Policy Statistics, Section on Statistics in Epidemiology
Tuesday, July 31, 2:00 pm–3:50 pm

Estimates of Intraclass Correlation from the Community Intervention Trial for Youth (CITY): An HIV Prevention Intervention for Young Men Who Have Sex with Men

❖Sherri L. Pals, Centers for Disease Control and Prevention, 1600 Clifton Road MS E45, Atlanta, GA 30333, *sfv3@cdc.gov*; Lillian S. Lin, Centers for Disease Control and Prevention; Carolyn Guenther-Grey, Centers for Disease Control and Prevention; Johnson Wayne, Centers for Disease Control and Prevention

Key Words: ICC, HIV, Sample Size, Group-randomized trials

Sample size estimation for group-randomized trials (GRTs) must take the intraclass correlation coefficient (ICC) into account, but few ICC estimates from HIV/STD prevention GRTs have been published. The CITY project evaluated an HIV risk reduction intervention targeting young men who have sex with men. We used two waves of CITY data to estimate ICCs for sex behavior variables, unadjusted and adjusted for covariates. The unadjusted ICC for any unprotected sex (last 90 days) ranged from about 0.005 to 0.012. With 200 members per community, these ICCs would result in variance at least twice as large as expected under individual randomization . Ignoring an ICC of this magnitude in planning a GRT could result in a substantially underpowered study. Our analyses provide estimates of ICCs from an HIV/STD prevention trial that may be useful to investigators designing future trials.

Effect of Multiple Imputation on the HIV Prevalence Estimate and Analysis of Risk Factors for HIV Infection: Guyana 2004

❖Maxine Denniston, Centers for Disease Control and Prevention, 4757 Banner Elk Drive, Stone Mountain, GA 30083, *mmd1@cdc.gov*; Nicole Seguy, Centers for Disease Control and Prevention

Key Words: multiple imputation, survey error, HIV risk factors, sample surveys, HIV prevalence estimation

In response to a report of high levels of HIV infection, a survey of miners in Guyana was fielded using 2-stage cluster sampling to estimate HIV prevalence and risks in this population. In 2006, results were presented from a sensitivity analysis of the effect on the HIV prevalence estimate of uncertainties in survey parameters due to challenges in survey implementation. This paper focuses on the effect of missing data on both the HIV prevalence estimate and risk factors for HIV infection. Multiple imputation was performed on HIV test and risk factor data using Sequential Regression Imputation. Of particular interest was the effect of imputing HIV test results for 35 miners who were interviewed but not tested. Results from this analysis will be compared to sensitivity analysis results. Independent risks for HIV infection will be determined and compared to those from the initial analysis.

On the Bias of Case Counts Adjusted for Reporting Delay

❖Ruiguang Song, Centers for Disease Control and Prevention, 1600 Clifton Road, Mailstop E48, Atlanta, GA 30333, *rsong@cdc.gov*

Key Words: Adjustment weight, Bias, Reporting delay, Trend analysis

To evaluate the trend in HIV diagnoses, the number of reported cases is adjusted for reporting delay to account for those cases recently diagnosed but not yet reported to the HIV case surveillance system. However, this adjustment has been observed to result in a positive bias in the estimated numbers of diagnosed AIDS cases. This bias can be significant in estimating the number of cases diagnosed near the ending date of the reporting period. In this talk, I will present the reason why the adjusted case count is positively biased and how large the bias could be. There are two ways to correct the bias: either by modifying the adjustment weights or by combining the adjustment weights using a double inverse approach. Simulation results on the performance of the proposed methods will be provided. Advantages and disadvantages of each method will be discussed.

A Comparison of Various Analytical Methods for U-Shape Distributed Random Variables: A Real-World Application from a Behavioral Intervention Trial

❖Terry Wang, Centers for Disease Control and Prevention, 1600 Clifton Rd NE MS E48, Atlanta, GA 30333, *dwang1@cdc.gov*; Torrance T. Stephens, Morehouse School of Medicine; Lillian S. Lin, Centers for Disease Control and Prevention

Key Words: cross section, proportional odds, survey, multivariate

HIV behavioral intervention trials target risky sexual behaviors for different at-risk populations. Various measures are used to assess changes in risky sexual behaviors, particularly condom use during sex, among the target populations. One construct typically used for analyses is the proportion of times a subject reports using a condom during a specific period. This random variable often follows a U-shaped distribution; standard linear regression procedures may not be valid. Other methods for operationalizing the outcome are also considered. We will present our findings from research studies based on the various operationalizations of condom use and will compare and contrast the various analytical approaches and their interpretations.

325 Statistical Methods for Studying Gene Expression and Transcriptional Regulation ● ✪

IMS, ENAR, Section on Physical and Engineering Sciences
Tuesday, July 31, 2:00 pm–3:50 pm

Clustering Time Course Gene Expression Data Using Nonparametric Hypothesis Testing

❖Haiyan Wang, Kansas State University, 108 E Dickens Hall, Department of Statistics, Manhattan, KS 66503, *hwanggo@gmail.com*; James Neill, Kansas State University; Forrest Miller, Kansas State University

Key Words: Clustering, Hypothesis testing, functional data, Multiple testing, High dimensional data, asymptotics

This paper presents an agglomerative clustering method for effectively detecting patterns and clusters in high dimensional functional data. The algorithm is based on two new nonparametric hypothesis tests applicable to curve screening and comparing groups of curves with heteroscedastic nonstationary observations over time. Flat curves are first screened and then two sets of curves are clustered into one group if they are not significantly different. The number of clusters are determined automatically by given significance levels. Our procedure takes into account both shape and magnitude of curves, is invariant under monotone transformation of data, and does not suffer from information loss due to the use of smoothing based procedures. Simulations show that the algorithm has low clustering error rates.

Studying the Relationship Between Coexpression and Coregulation of Protein Complex Using ChIP-Chip Data

❖Ching-Ti Liu, Yale University, 94 Howe St APT 4, New Haven, CT 06511, *ching-ti.liu@yale.edu*; Ker-Chau Li, University of California, Los Angeles

Key Words: protein complex, microarray

Characterization of protein-protein interaction has become a key step in understanding the cellular mechanism because most biological processes cannot be carried out without protein-protein interaction. This presentation will focus on a set of well-validated protein complex data from MIPS. We will discuss the way of incorporating the ChIP-on-Chip data for protein complexes analysis.

Protein Multiple Alignment Using Primary and Secondary Structure Information

❖Nak-Kyeong Kim, National Institutes of Health, NCBI, 8600 Rockville Pike, Bldg 38A Rm 6N611D, Rockville, MD 20894, *kimnak@ncbi.nlm.nih.gov*; Jun Xie, Purdue University

Key Words: Gibbs sampling, likelihood function, protein sequence motifs, secondary structures, segment overlap

Identifying common local segments, also called motifs, in multiple protein sequences plays an important role for establishing homology between proteins. Homology is easy to establish when sequences are similar. However, for distant proteins, it is much more difficult to align motifs that are not similar in sequences but still share common structures or functions. This is a first attempt to align multiple protein sequences using both primary and secondary structure information. A new sequence model is proposed so that the model assigns high probabilities not only to motifs that contain conserved amino acids but also to motifs that present common secondary structures. The proposed method is tested in a structural alignment database BAliBASE. Test results show that information brought by the predicted secondary structures greatly improves motif identification.

Detection of eQTL Modules Mediated by Activity Levels of Transcription Factors

❖Wei Sun, University of California, Los Angeles, 11130 Rose Ave Apt 403, Los Angeles, CA 90034, *sunwei@ucla.edu*; Ker-Chau Li, University of California, Los Angeles; Tianwei Yu, Emory University

Key Words: eQL, gene expression, transcription factor, causal analysis, TF activity

Studies of gene expression quantitative trait loci (eQTL) have shown the existence of eQTL hotspot. Two questions of great interest about eQTL hot-spots arise: (1) Which gene is the quantitative trait gene (QTG)? (2) How does a QTG affect expression levels of many genes linked to it? The first question can be answered by biological evidence or statistical methods. The second question is harder to address. One simple situation is that the QTG encodes a transcription factor (TF), which regulates the expression of genes linked to it. However, previous results have shown that TFs are not over-represented in the eQTL hotspots. We consider the scenario that the propagation of genetic perturbation from a QTG to other linked genes is mediated by the TF activity. We develop a procedure to detect the eQTL modules (eQTL hot-spots together with linked genes) that are compatible with this scenario.

Variable Selection Using Single Index Models for Motif Discovery

❖Wenxuan Zhong, Harvard University, One Oxford St, Science Center 701, Cambridge, MA 02138, *wzhong@fas.harvard.edu*

Key Words: variable selection, single index model, dimension reduction, BIC, stepwise regression, motif discovery

Information for regulating a gene's transcription is contained in the conserved patterns (motifs) on the upstream/downstream DNA sequence (promoter region) close to the target gene. By combining the information contained in both gene expression measurements and genes' promoter sequences, I proposed a novel procedure for identifying functional active motifs under certain stimuli. A nonlinear regression model, single index model, was used to associate promoter sequence information of a gene and its mRNA expression measurements. In this talk, I will describe computational efficient variable selection procedures and criteria, which were developed for the single index model. I will also demonstrate the advantage of these methods both theoretically and empirically. Compared with existing methods, our proposed procedures can greatly improve variable sensitivities and specificities.

326 Applications of Bayesian Inference in Wildlife Ecology ● ✪

Section on Bayesian Statistical Science, ENAR, Section on Statistics and the Environment

Tuesday, July 31, 2:00 pm–3:50 pm

A Hierarchical Bayesian Model for Stochastic Metapopulation Viability in the Presence of Translocation

❖Christopher K. Wikle, University of Missouri-Columbia, 146 Middlebush, Columbia, MO 65211, *wiklec@missouri.edu*; Sammuel F. Bussman, University of Missouri-Columbia; Robert L. Schooley, University of Illinois

Key Words: ecology, metapopulation, translocation, population viability

Active management of a population through translocation of individuals from a thriving subpopulation to less healthy subpopulations has proven to be a successful approach for improving metapopulation viability. However, surprisingly little empirical work has been done in which metapopulation viability models accounting for translocation have been considered in the presence of uncertainty. We consider such a model in the Bayesian hierarchical paradigm. The model is applied to the endangered red-cockaded woodpecker, using data from military lands in the U.S. Southeast.

Discrete Representations of Ecological Process Models in the Presence of Uncertainty and Data

❖Mevin Hooten, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, *mevin.hooten@usu.edu*; Amanda R. Cangelosi, Utah State University

Key Words: differential equations, hierarchical models, population biology

Hierarchical models are being increasingly utilized to implement deterministic physical process models in an inverse fashion; that is, in settings where observed data are utilized to estimate model parameters in the presence of uncertainty. The hierarchical statistical approach generally incorporates a discretized version (e.g., difference equation) of the dynamical process (e.g., differential equation) and allows for the rigorous estimation of latent state variables, model parameters, and variance components. Despite the success of such implementations, little attention has been paid to the precision of the approximated deterministic model itself. Here we implement and compare varying levels of approximated dynamical processes in a hierarchical setting and illustrate by modeling a true ecological process.

Continuous-Time Modeling of Marine Mammal Telemetry Data

❖Devin Johnson, National Marine Mammal Laboratory, 7600 Sandpoint Way NE, Seattle, WA 98012, *devin.johnson@noaa.gov*; John Durbin, National Marine Mammal Laboratory; Mary-Anne Lea, National Marine Mammal Laboratory; Joshua London, National Marine Mammal Laboratory

Key Words: ARGOS, Continuous-time, Animal movement, Stochastic process, Marine Mammals, Telemetry data

Marine mammal telemetry data is obtained by determining the location of an animal in space at several points in time. The observed locations are function of two important factors, true location and measurement error. In order to handle the fact that most telemetry data is collected opportunistically on an irregular basis we consider a movement model stochastic process in continuous time. Using this approach likelihood can be formed using the raw data instead of aggregated, thinned, or interpolated data. Bayesian methods are explored for making inference on animal locations as well as other movement quantities of interest, such as travel speed. The Bayesian inference paradigm allows for inclusion of parameter uncertainty in location estimation, as well as, propagation of uncertainty through nonlinear relationships in the movement parameters of interest.

Estimation of Carcass Recovery Rates for Florida Manatees Using MCMC

❖Christopher Fonnesbeck, Fish & Wildlife Research Institute, FWC, 1250 Woodland Avenue SE, Atlanta, GA 30316, *chris.fonnesbeck@myfwc.com*; Carol Sanders-Reed,

Key Words: mortality, bayesian, MCMC, manatee, recovery

Adult mortality has been cited as an important threat to the recovery of the endangered Florida manatee (Trichechus manatus latirostris). In an effort to assess manatee mortality, the Florida Fish & Wildlife Conservation Commission retrieves all carcasses reported throughout the state by the public. However, the reporting and recovery rate of carcasses, and the dynamics thereof, are unknown. This prevents using such information for inference regarding the corresponding rate of mortality. We apply recovery models (e.g., Brownie et al. 1985), typically applied to band recoveries of harvested waterfowl, to the problem of manatee carcass recovery estimation. Data from recovered manatees previously marked with PIT tags, along with relevant covariates, are used to inform the estimation model. We express the model in a Bayesian context and estimate parameters using Markov chain Monte Carlo.

Modeling Unobserved Sources of Heterogeneity in Animal Abundance Using a Dirichlet Process Prior

❖Robert Dorazio, U.S. Geological Survey/University of Florida, Dept of Statistics, Gainesville, FL 32611-0339, *bdorazio@usgs.gov*; Bhramar Mukherjee, University of Michigan; Li Zhang, The Cleveland Clinic; Malay Ghosh, University of Florida

Key Words: abundance estimation, detection heterogeneity, Dirichlet process, Bayesian nonparametric

In surveys of natural animal populations, a sampling protocol is often spatially replicated to collect a representative sample of the population. In these surveys differences in abundance of animals among sample locations may induce spatial heterogeneity in the counts associated with a particular sampling protocol. For some species, the sources of heterogeneity in abundance may be unknown or immeasurable, leading one to specify the variation in abundance among sample locations stochastically. However, choosing a parametric model for the distribution of unmeasured heterogeneity is potentially subject to error and can have profound effects on predictions of abundance at unsampled locations. We develop a robust alternative approach wherein a Dirichlet process prior is assumed for the distribution of latent abundances.

327 Nonparametric and Robust Methods for Genetic Data Analysis ● ✪

Section on Nonparametric Statistics, WNAR

Tuesday, July 31, 2:00 pm–3:50 pm

Empirical Likelihood-Based Inferences in a Complex Genetic Mixture Model

❖Jing Qin, National Institutes of Health, 6700A Rockledge Drive, Room 5131, Bethesda, MD 20817, *jingqin@niaid.nih.gov*

Key Words: Density ratio model, Empirical likelihood, Mixture model, Semiparametric method

In genetic experiments, data are often collected from complex mixtures of distribution functions with known mixing proportions. Previous approaches to estimation involve parametrizing the distributions. In this paper we explore a semiparametric alternative model used by Anderson (1979), in which the log ratio of probability (density) functions from components is linear in the observations. Under this semiparametric setup, we show that the multivariate mixture model is identifiable. However, the univariate semiparametric marginal model can not identify the underlying parameters. Profile empirical likelihood technique is applied for estimation and test problems. It is shown that the maximum empirical likelihood estimate has an asymptotic normal distribution. The empirical likelihood ratio statistic behaves like a chi-square variable. As a by product, the underlying distribution

A Method Toward Mapping of Common Traits

❖Shaw-Hwa Lo, Columbia University, Department of Statistics, 1255 Amsterdan Ave. 10th floor, New York, NY 10027, *slo@stat.columbia.edu*; Tian Zheng, Columbia University

Key Words: genetic mapping, common traits, interactive effects

The mapping of complex/common traits is an important yet difficult research area. Many common human disorders are believed to be "complex" or multifactorial, meaning that they cannot be attributed to alleles of a single gene or one risk factor. Mapping methodologies that simultaneously inspect disjoint loci are crucial for the success of genes mappings. We consider an alternative method (Lo and Zheng 2002, 2004) proposed recently, using family-trio data and several disease models, followed by a demonstration of applying this method to IBD data. The outcomes of this practice suggested the potentials of this approach in drawing substantially more information, leading to novel scientific findings. The main idea applies more generally than to special genetic problems. If time permits, a general mathematical framework will be introduced.

Two-Stage Tests for Association on the X Chromosome

❖Jungnam Joo, National Heart, Lung, and Blood Institute, NIH, 6707 Democracy Blvd, Suite 105, Bethesda, MD 20892-5478, *jooj@nhlbi.nih.gov*

Key Words: association, Two-stage test, Genome-wide association study, significance level

Testing association for markers on the X chromosome possesses a unique feature due to the fact that males carry only one copy of X chromosome. Compared to the extensive research that has been done for markers on the autosomal chromosomes, few work has been done for markers on the X chromosome. In my talk at the last JSM, I presented several combinations of two test statistics based on male and female samples separately. In this talk, I will present several two-stage tests where males are used in the first stage and females are used in the second stage. The choices of significance levels for the two stages are discussed. Simulation studies are conducted to examine the properties of the proposed method, and the results are applied to data from a genome-wide association study for age-related macular degeneration (AMD).

Genetic Model Selection in Two-Stage Analysis for Case-Control Association Studies

❖Hon Keung (Tony) Ng, Southern Methodist University, Department of Statistical Science, Southern Methodist University, Dallas, TX 75275, *ngh@mail.smu.edu*; Gang Zheng, National Heart, Lung, and Blood Institute

Key Words: Efficiency robustness, Genetic models, Genome-wide association, Hardy-Weinberg disequilibrium, Two-stage analysis

The genotype-based Cochran-Armitage trend test (CATT) has been suggested for testing association between a marker and a disease in case-control studies. When the underlying genetic model is known, the CATT with the recessive, additive, and dominant scores is used. However, the genetic model is often unknown. Therefore, robust tests, such as the CATT with the additive score, the maximum efficiency robust test (MERT) and the maximum tests (MAX) are used in practice. We propose a two-stage analysis: In the first stage, we use the difference of Hardy-Weinberg disequilibrium coefficients between the cases and controls for model selection; in the second stage, an optimal CATT, corresponding to the selected model, is used for testing association. Simulation studies show that this approach is more preferable to the CATT with the additive score and the MERT, and it is more powerful than the MAX.

A New Hypothesis-Testing Approach When the Nuisance Parameter Is Only Present Under the Alternative Hypothesis

❖Olivier Thas, Ghent University, Coupure Links 653, Gent, 9000 Belgium, *olivier.thas@ugent.be*; Gang Zheng, National Heart, Lung, and Blood Institute

Key Words: hypothesis testing, nuisance parameters, nonparametric test, goodness-of-fit, empirical process, genetics

We present a new class of statistical tests. The statistic is defined as the length of an excursion set (LESS) of an empirical process. Examples of testing situations where the test may be useful include goodness-of-fit, linkage analysis, and testing a hypotheses when a nuisance parameter is present only under the alternative. We give the asymptotic distributions under both the null hypothesis and under contiguous alternatives. By defining an excursion as an exceedance of an upper or lower boundary function, these functions may be chosen to get good power. In a simulation study, we empirically investigate the behavior of the tests, and we illustrate its application to the examples mentioned earlier.

328 Statistical Methods for Graphs and Networks

Section on Statistical Computing, Section on Statistical Graphics

Tuesday, July 31, 2:00 pm–3:50 pm

Bayesian Self-Modeling Warping Regression

❖Donatello Telesca, University of Washington, 1816 Boylston Ave 202, Seattle, WA 98122, *telesca@stat.washington.edu*; Lurdes Inoue, University of Washington

Key Words: Hierarchical Models, Curve Registration, MCMC, Microarray Data, Network Inference

Functional data often exhibit a common shape but also variations in amplitude and phase across curves. The analysis often proceed by synchronization of the data through curve registration. In this paper, we propose a Bayesian Hierarchical model for curve registration. Our hierarchical model provides a formal account of amplitude and phase variability while borrowing strength from the data across curves in the estimation of the model parameters. We discuss extensions of the model by utilizing penalized B—splines in the representation of the shape and time—transformation functions, and by allowing random image sets in the time transformation. We discuss applications of our model to simulated data as well as to two datasets. In particular, we illustrate using our model in a non–standard analysis aimed at investigating regulatory network in time course microarray data.

What Is a 'Random Network'?

❖David Hunter, The Pennsylvania State University, 310 Thomas Building, University Park, PA 16802, *dhunter@stat.psu.edu*

Key Words: Networks, Social networks

This talk gives a brief survey of some existing ideas about stochastic models for networks. We will discuss a couple methods of randomly generating networks, including methods for generating networks with a specified degree distribution. The talk is also a bit of a plea for statisticians to become more involved in social networks research, as we point out some misapplications of statistical ideas that exist in the current literature.

Collective Inference for Network-Based Marketing

❖Shawndra Hill, University of Pennsylvania, 3730 Walnut Street, Suite 500, Philadelphia, PA 19104, *shawndra@wharton.upenn.edu*

Key Words: collective inference, social network analysis, graphical models, viral marketing, network-based marketing

Data on large consumer networks are becoming increasingly available. For example, e-commerce and telecommunications firms are in the business of collecting data on their large proprietary consumer networks and portal sites such as MySpace and Facebook provide publicly available data on links between networked consumers to members of their network. As these consumer network data continue to become larger, firms will need better techniques to use network features for business decision-making. This research adapts collective inference models applied previously in other domains such as image processing and web page classification, to identify likely adopters of a large telecommunication firm's new technology service. We show that our collective inference method, a global approach, improves consumer ranking over existing methods that use only local network features.

Transitivity in Weighted Graphs: Effects on the Topology of Knowledge and Social Networks

❖ Tiago Simas, Indiana University, 1900 E 10th St, Bloomington, IN 47406, *tdesimas@indiana.edu*; Bharat Dravid, Indiana University; Luis Rocha, Indiana University

Key Words: Weighted graphs, Transitivity, Transitive Closure, Scale Free Networks, Complex Networks

There are different ways to calculate transitivity in weighted graphs. Most real weighted graphs are semimetric (i.e., symmetric, reflexive, but not transitive). Modelers typically embed these graphs in a metric space using the Dijkstra algorithm, which is in effect a metric closure that causes some distortion of the topology of graphs. We relate metric properties of distance graphs with the transitive properties of proximity graphs. The isomorphism between these graphs is formulated via t-norm/t-conorm functions. We show (1) the most intuitive t-norm pair (max, min) leads to an ultra-metric closure with the loss of scale-free structure. (2) The metric closure leads to a t-norm pair with poor axiomatics. A sounder transitive closure for complex networks is given by Hamacher's t-norm pair, which is shown to better preserve the scale-free structure of real networks.

329 Studies with Multiple Biomarkers ●

Section on Statistics in Epidemiology, Biometrics Section, ENAR

Tuesday, July 31, 2:00 pm–3:50 pm

Diagnostic and Analytical Accuracies of Qualitative Multiplex Tests

❖ Marina Kondratovich, Food and Drug Administration, CDRH, HFZ550 1350 Piccard Dr, Rockville, MD 20850, *marina.kondratovich@fda.hhs.gov*

Key Words: Diagnostic Device, Multiplex Test, Analytical Accuracy, Limit of Detection, Diagnostic Accuracy, Sensitivity and Specificity

With increased frequency, medical tests are being designed as measuring devices of multiple (at least 2) measurands tested simultaneously in a single patient sample (multiplex tests). Multiplex tests generate multiple outputs on a single patient sample. We consider multiplex assays with a qualitative output for each measurand. When test reactions for multiple measurands occur simultaneously, analytical performance will depend on potentially complex interactions. Issues related to analytical performance of multiplex assays as analytical accuracy (measures of agreement of how well a multiplex assay detects presence or absence of each measurand separately and combination of measurands), limits of detections, and precision are considered. We present also issues related to reporting diagnostic accuracy (measures of agreement of outcomes of multiplex assay with clinical target condition).

Interpretation of Multimarker Scores

❖ Ursula Garczarek, Roche Diagnostics GmbH, Nonnenwald 2, Penzberg, 82377 Germany, *ursula.garczarek@roche.com*

Key Words: multi-marker panel, probability score

Medical diagnosis can be based on various clinical data (e.g., from imaging techniques, physical exams, and laboratory). Physicians combine these to arrive at a diagnostic conclusion. From that perspective multi-marker panels for diagnosis are intended to pre-aggregate some of the clinical data in an objective manner. Often the diagnostic result of some marker-panel is expressed by some score that is larger the more likely the medical condition at question is. As there is no obvious choice of scale for the score, physicians need advice on how to interpret these results. It is tempting to advocate a straight probability interpretation. Yet, what type of probability is appropriate? How should the validity of the probability claims be demonstrated? We will present some strategies to scale scores along with benefits and risks in their interpretation and performance evaluation.

Multivariate Assays Reporting a Composite Score

❖ Estelle Russek-Cohen, Food and Drug Administration, 1350 Piccard Dr, CDRH, Rockville, MD 20850, *estelle.russek-cohen@fda.hhs.gov*; Mary Pastel, Food and Drug Administration; Zivana Tezek, Food and Drug Administration

Key Words: In vitro diagnostics, medical devices, microarrays

With the rise of genomic and proteomic technologies, one would expect to see medical devices that combine the results of multiple probes, analytes or biomarkers into a single result or score that is then interpreted. These assays need to be sufficiently robust that they can be performed over long time periods and in multiple labs and still be reproducible. This talk will include a discussion of how performance in these in vitro devices could be measured, whether a device is associated with a diagnostic or prognostic intended use. The role of an external validation dataset will also be covered.

Principal Component Analysis for Data Reduction of Gait Analysis for Knee Prosthesis Research

❖Jeffrey Murphy, DePuy Orthopaedics, 700 Orthopaedic Drive, Warsaw, IN 46582, *jmurphy7@dpyus.jnj.com*; Luke Aram, DePuy Orthopaedics; Jordan Lee, DePuy Orthopaedics; Travis Bennett, DePuy Orthopaedics; Catherine Truxillo, SAS Institute Inc.; Paul Rullkoetter, University of Denver

Key Words: principal components analysis, knee prosthesis design, DOE, T-square, scree plot, Gait analysis

Principal components analysis is useful in reducing voluminous gait data into manageable amounts while retaining much of the original information. Thirty factors in knee prosthesis design were set at two levels in a resolution IV fractional factorial DOE. Finite element analysis virtually built each run. Fifteen responses were measured. Principal components (PCs) were retained using the scree plot elbow method. Varimax orthogonal rotation and waveform plots were used when multiple PCs were retained. Influential runs were identified using T-square statistics. Between one and four PCs were retained for each response. Sagittal major radius and femoral internal/external rotation were the two most significant factors. Level setting combinations were carefully reviewed on influential runs. This was successful in reducing gait data into the critical few factors vital to knee prosthesis design.

Utilizing Multiple Biomarkers To Assess Myocardial Function in a Dose-Finding Study: Determining the Optimal Dose by 'Scoring' the Outcomes

❖ Cynthia L. Green, Duke Clinical Research Institute, 2400 Pratt St, Durham, NC 27705, *green054@mc.duke.edu*; Kathleen Trollinger, Duke Clinical Research Institute; Saling Huang, KAI Pharmaceuticals; Matthew Roe, Duke Clinical Research Institute

Key Words: surrogate, biomarkers

Biomarkers are often chosen as primary endpoints in Phase I/II trials. Biomarkers are intended to simulate the effect of the treatment on clinical outcome by providing a surrogate known to have strong effects on measures of the biological process. Multiple biomarkers have been studied which correlate to myocardial function, including TIMI flow, enzymes, sestamibi scans, ejection fraction and ST resolution. The DELTA-MI study evaluated the safety and efficacy of KAI-9803 in 154 patients with acute MI. Subjects were randomized 2:1 to KAI-9803 or placebo into one of four dosing groups (0.05, 0.5, 1.25 and 5.0 mg). Multiple biomarkers of myocardial function were analyzed to determine the optimal dosing level. Our goal was to provide a statistical model to determine the optimal dose by "scoring" the direction and amount of treatment benefit for each biomarker within each dose group.

330 Multiple Frame Survey Design and Inference

Section on Survey Research Methods

Tuesday, July 31, 2:00 pm–3:50 pm

A Case Study in Dual-Frame Estimation Methods

❖ Keith Rust, Westat, 1650 Research Blvd, Rockville, MD 20850, *keithrust@westat.com*; Inho Park, Westat; Marianne Winglee, Westat; Gary Shapiro, Westat

Key Words: Multi-stage sampling, Sampling weights, Rare populations, Domain estimation, Single-frame (SF) estimator, Pseudo-maximum likelihood (PML) estimator

In the National Incidence Study of Child Abuse and Neglect (NIS), the survey uses a multistage and multiframe design to broaden coverage of possible reporting sources in order to produce more accurate and efficient estimates. Within sampled counties the NIS collects lists of maltreated children investigated by Child Protective Services (CPS) agencies and those seen by samples of professionals in agencies such as police departments, hospitals, and schools. While the list frame of children investigated is known for CPS agencies, only sampled data are available for the other agency categories. As a result duplication of cases across agencies cannot be fully identified. We present a modified single-frame estimator designed to accommodate this situation. This estimator is compared with the classic single-frame estimator and the pseudo-maximum likelihood estimator in a simulation study.

A Strategy for Estimating the Number of Minority-Operated Farms

❖ Matthew J. Fetter, USDA National Agricultural Statistics Service, 1400 Independence Ave SW, RM 6436A, Washington, DC 20250, *matt_fetter@nass.usda.gov*; Phillip S. Kott, USDA National Agricultural Statistics Service

Key Words: area segment, nearly unbiased, variance estimation, weight adjustment, selection probability, supplemental sample

The National Agricultural Statistics Service (NASS) uses a national stratified sample of area segments to measure production agriculture. It has supplemented this sample to estimate the total number of 2007 farms in specific categories, in particular the number farms operated by minorities. To meet a national accuracy criterion for the number of Asian-operated farms yet prevent unreasonably large sample sizes in some strata in California, NASS has designated specific segments for potential supplemental sampling based on Decennial Census information. We will show how this was accomplished without sacrificing the randomization consistency of the Asian estimators or increasing the anticipated mean squared errors of other estimators.

Linearization Variance Estimators for Dual Frame Survey Data

❖ Abdellatif Demnati, Statistics Canada, Social Survey Methods Division, Coats Building 15th Floor G, Ottawa, ON K1A0T6 Canada, *abdellatif.demnati@statcan.ca*; J. N. K. Rao, Carleton University; Mike A. Hidiroglou, Statistics Canada; Jean-Louis Tambay, Statistics Canada

In sampling from a single complete frame, Taylor linearization is often used to obtain variance estimators of calibration estimators of totals and nonlinear finite population parameter. Demnati and Rao (2004) proposed a new approach to deriving Taylor linearization variance estimators that leads directly to a unique variance estimator that satisfies some desirable properties for general designs. With increase in the number of household surveys, the cost of personnel interviewing has increased significantly. As a result, new surveys are often conducted using dual frames: a complete area frame and an incomplete telephone frame. This paper first describes some dual frame estimators based on multiple weight adjustments. The Demnati-Rao method is then applied to take into account such multiple weight adjustment for variance estimation.

Designing a Dual-Frame Survey for Breeding Bald Eagle Pairs in the Lower 48 States

❖ Mark C. Otto, U.S. Fish & Wildlife Service, 2302 Lackawanna Street, Adelphi, MD 20783, *mark_otto@fws.gov*; John R. Sauer, U.S. Geological Survey

The USGS and US Fish and Wildlife Service are designing a dual frame survey of breeding bald eagle pairs in the lower 48 states following Haines and Pollock, 1998. The list frame is made up of the combined state nests lists, and the area frame is a grid of 10 by 10 km2 plots. The initial state bird conservation region strata were collapsed sequentially by combining adjacent strata with similar densities. Cost functions of aerial miles were made up from simulating samples of different sizes from the nest lists. Optimal sampling allocations between the list and area frames were determined over different list coverages. Efficiencies for simultaneously sampling the area then list frames will also be discussed.

331 Joint Analysis of Longitudinal and Survival Data ✪

WNAR, Biometrics Section, ENAR, Section on Risk Analysis, Section on Statistics in Epidemiology

Tuesday, July 31, 2:00 pm–3:50 pm

Joint Modeling of Recurrent Event and Event Marker

❖Jianwen Cai, The University of North Carolina at Chapel Hill, Department of Biostatistics, CB7420, Chapel Hill, NC 27599-7420, *cai@bios.unc.edu*; Donglin Zeng, The University of North Carolina at Chapel Hill

Key Words: recurent event, event marker, joint modeling, estimating equations, semiparametric models

In many biomedical studies with recurrent event, some markers can only be measured when the event happens. For example, in a study for hospitalization, medical cost can only incur when the patient is hospitalized. It is of interest to study the effect of some risk factors on both rate of hospitalization and the cost associated with the hospitalization. In this talk, we present simultaneous modeling of both recurrent event and the event marker using the observed covariates. Our semiparametric models assume marginal relationship between covariates and both recurrent event and event marker and it naturally follows from general random effect models. Estimating equations are constructed to derive the estimates for the parameters in the model. The asymptotic properties are established and finite sample properties are studied. The proposed method is applied to a real data set for illustration.

Robust Inference for Joint Modeling of Longitudinal Measurements and Competing Risks Failure Time Data

❖ Ning Li, University of California, Los Angeles, 3281 S Sepulveda Blvd, Apt 307, Los Angeles, CA 90034, *ningli@ucla.edu*; Gang Li, University of California, Los Angeles; Robert Elashoff, University of California, Los Angeles

Key Words: Cause-specific hazard, Competing risks, EM algorithm, Joint modeling, Longitudinal data, Robust inference

We propose a robust joint model for analysis of longitudinal measurements and competing risks failure time data in the presence of substantial outlying longitudinal observations during follow-up. Our model consists of a linear mixed effects sub-model for the longitudinal outcome and a proportional cause-specific hazards frailty sub-model for the competing risks data, linked together by some latent random effects. We study the maximum likelihood estimates of the parameters by an EM algorithm and estimate their standard errors using a profile likelihood method. The robustness of normality assumption for the random effects in the linear mixed sub-model is also investigated.

Joint Analysis of Longitudinal Measurements and Competing Risks Failure Time Data: A Bayesian Approach

❖ Wenhua Hu, University of California, Los Angeles, 3232 Sawtelle Blvd, Apt 104, Los Angeles, CA 90066, *hwh@ucla.edu*; Gang Li, University of California, Los Angeles

Key Words: Competing risks, joint longitudinal and survival model, informative censoring, Bayesian Modeling

This paper proposes a Bayesian method for joint modeling of longitudinal and competing risks survival data. The outcomes are modeled by observed covariates as well as by some unobserved factors which introduce latent association between the longitudinal model and competing risks model. This approach is useful in many applications since it allows us to analyze the longitudinal outcome with non-ignorable missing data induced by multiple types of events, to analyze longitudinal and survival data with informative censoring for the key event by treating it as a competing risk, and to model both endpoints simultaneously given both observed covariates and unobserved factors. We present an analysis for a data set from a clinical trial of the scleroderma lung disease. The performance of our joint modeling approach is evaluated through simulations.

Joint Analysis of Repeated State Transition in Activities of Daily Living and Death in Longitudinal Aging Study

❖ Haiqun Lin, Yale University, *haiqun.lin@yale.edu*; Zhenchao Guo, Yale University; Peter Peduzzi, Yale University; Thomas Gill, Yale University; Heather G. Allore, Yale University

Key Words: aging study, dependent censoring, latent trait, longitudinal data, multistate transition

A general multistate transition model is proposed for repeated episodes of multiple states representing different health status. A semiparametric model with transition-specific nonparametric baseline intensities and state-specific latent traits is adopted. A latent trait captures individual tendency of repeated sojourn in a state and accounts for within-state-within-subject correlation. Between-state correlation is accounted for by the covariance of the latent traits. The factor loadings of a latent trait accommodate the dependence of transitions to different competing states from a same state. The method is illustrated by repeated transitions between independence and disability of activities of daily living with death as an absorbing state in a longitudinal aging study. A simulation study is performed to study the appropriateness of the asymptotic estimates of the standard errors.

332 Applications of Bayesian Statistical Method ● ✪

Biometrics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Estimating the Survivor Average Causal Effect in Nonrandomized Studies Where Treatment Assignment Is Missing for Some Patients

❖ Abbie Stokes-Riner, University of Rochester, Department of Biostatistics, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *abbie_stokesriner@urmc.rochester.edu*; Jason Roy, Geisinger Health System; Sally W. Thurston, University of Rochester

Key Words: causal effect, truncation due to death, missing data, Bayesian inference, posttreatment variable

Evaluating the causal effect of treatment can be complicated when the outcome measurement is truncated due to a post treatment event. For example, obstetricians are interested in the effect of early epidural use on the duration of second stage labor. However, this measurement will not be available if the patient had a cesarean section delivery (C-section). The situation is further complicated when the treatment for some patients is unknown. We take a fully Bayesian approach to (1) adjust for confounding due to non-randomization, and (2) impute missing treatment assignments using auxiliary information. A sensitivity analysis is proposed to account for assumptions whose validity cannot be assessed from the data. The method is applied to obstetric data, to identify the effect of early epidural use on duration of second stage labor.

Bayesian Modeling of Multivariate Binary Response with a Change-Point for Applications to Adolescent Offender Registry

❖ Dipankar Bandyopadhyay, Medical University of South Carolina, 135 Cannon Street Suite 303, DB2E MUSC, Charleston, SC 29425, *bandyopd@musc.edu*; Debajyoti Sinha, Medical University of South; Stuart R. Lipsitz, Brigham and Women's Hospital; Elizabeth Letourneau, Medical University of South Carolina

Key Words: binary, bridge, changepoint, logit, multivariate

We develop a model for multivariate binary data using random effects logistic regression model incorporating a change-point. To preserve marginal logit structure (integrating out random effects) for interpretable co-variate effects, we model the random effects using a bridge density and compare it with normal random effects that doesn't preserve the logit link marginally. We analyze the data from South Carolina adolescent sex repeat-offenders to investigate the effects of different factors and presence of any change-point following enactment of a new set of mandatory sentencing laws (occurring between 1992–1996) regarding the solicitors' decision and judges' adjudication. A discrete prior on the change-point is elucidated to reflect the expert's opinion on possible year of change in judges/solicitors opinions and posteriori conclusion on changepoint occurrence between 1992–1996.

Bayesian Genome-Wide Screening via Pseudo Perfect Sampling for Quantitative Trait Loci Problems

❖ Cheongeun Oh, University of Medicine and Dentistry of New Jersey, 92 Mercer St Apt#2, Jersey City, NJ 07302, *ohch@umdnj.edu*

Key Words: Model selection, quantitative trait, MCMC

With recent interest being focused on genome scans using a large number of marker loci, the common approaches of existing statistical methods produce often the inconsistent results. The methodologies using Bayesian model selection via Markov chain Monte Carlo sampling schemes have been developed to remedy the pitfall of classical methods, despite biases that might be introduced in the inferences. In this paper, we propose robust and effective Bayesian genome-wide screening methods of simultaneously identifying loci related to quantitative trait and their possible interaction effects over the entire genome. Our methods are based on Bayesian model selection, and for posterior inferences, we utilize the emerging simulation technique, Perfect sampling algorithm to alleviate any concerns over convergence and sampling mixing, thus leading to a powerful and robust performance.

Ridge Regression To Accommodate LD in WGA Studies

❖Nathalie Malo, University of California, San Diego, 3855 Health Science Drive, San Diego, CA 92093, *nmalo@ucsd.edu*; Jennifer Wessel, University of California, San Diego; Nicholas J. Schork, University of California, San Diego

Key Words: association analysis, epistasis, genetic variation, genome-wide scans, target SNPs

The use of whole genome association (WGA) studies for the identification of genes and genetic variations that influence common, complex diseases such as hypertension, cancer, and depression will continue to grow as cost-effective high-throughput genotyping technologies are developed. As a result, appropriately flexible yet robust data analysis strategies for analyzing WGA data will be essential. We emphasize the need to accommodate phenomena such as linkage disequilibrium via simple extensions of traditional regression models. We describe the use of regression analysis models for WGA that are very intuitive and flexible. We propose the use of ridge regression, a special case of Bayesian regression, to account for correlation. We showcase the utility of the method on previously published WGA data. We also consider limitations of the proposed approach as well as areas for further research.

Small Sample Conditional Inference for the Cox Model

❖Juan Zhang, Rutgers University, Dept of Statistics and Biostatistics, 501 Hill Center, Piscataway, NJ 08855, *janezh@stat.rutgers.edu*; John E. Kolassa, Rutgers University

Key Words: Saddlepoint approximation, Cox model, Conditional inference, Modified signed likelihood ratio statistic, Priors, Partial differential equation

We apply saddlepoint approximation to conditional inference under the Cox model. With the presence of nuisance parameters, inference about the parameter of interest may be based on the signed root of the likelihood ratio statistic. DiCiccio and Martin (1993) proposed an alternative quantity to the signed root of the likelihood ratio statistic, which involves the Bayesian approach. The prior function of the parameters can be obtained from a partial differential equation. For the Cox model, finding the prior function is not trivial. We provide a way to solve the partial differential equation, apply it to the Cox model, and perform numerical studies in small sample size settings.

A Bayesian Analysis of Spatially Correlated Binary Data with Applications in Dental Research

❖Yanwei Zhang, Michigan State University, Division of Biostatistics, Department of Epidemiology, East Lansing, MI 48824, *zhangy@stt.msu.edu*; David Todem, Michigan State University; R. V. Ramamoorthi, Michigan State University

Key Words: Spatial Correlation, Multilevel Nested Co-dependency Structure, Finite Mixture Model, Undirected Graphical Model, Bayesian Hypothesis Testing

Analysis of dental caries is traditionally based on DMFS and DMFT scores. Although this approach has aided our understanding of the pattern of dental caries, there are still some fundamental unanswered questions. As an example, it is believed among dentists that there are spatial symmetries in the mouth with respect to caries, but this has never been shown in a statistical sense. An answer to this question requires the analysis to be performed at the tooth level. In this paper, we formulate a finite Bayesian mixture model coupled with an undirected graphical approach to accommodate the nested co-dependency structure of dental data. The spatial symmetry of the development of dental caries in the mouth is assessed using the Bayesian principle. Data from a cross-sectional survey are used to illustrate the method.

Hierarchical Mixed Membership Models for Honey Bee Genomes

❖ Tanzy Love, Carnegie Mellon University, Department of Statistics, Pittsburgh, PA 15213, *tanzy@cmu.edu*

Key Words: mixed membership, soft clustering, honey bees, bayesian, dirichlet process prior

In the New World, all existing Apis mellifera (honey bees) are introduced from existing (different) populations in the Old World. One effect of this is the prevalence of European honey bees in the north and African “killer” honey bees in south (these frighten some by moving). Using 1136 SNPs genotyped in 341 individuals, Whitfield et al. (2006) characterized both native and introduced populations of honey bees over time. In their paper, they clustered native genomes into four genetic classes, and then compared introduced genomes to these clusters to measure the change in the honey bee gene pools in the Americas over time. We propose a mixed membership model for honey bee genomes where any particular bee’s genome is a mixture of several pure types. Using a Dirichlet process prior, we can estimate the true number of pure types along with the mixture in each of the observed individuals.

333 Robust Modeling Strategy ● ✪

Biometrics Section, Section on Nonparametric Statistics

Tuesday, July 31, 2:00 pm–3:50 pm

Variability Study of a Robust Estimator of Mutation Rate in Fluctuation Analysis

❖ Xiaowei Wu, Rice University, 6100 Main St MS138, Houston, TX 77005, *xwwu@stat.rice.edu*; Erin Strome, Texas Children’s Cancer Center; Sharon Plon, Texas Children’s Cancer Center; Marek Kimmel, Rice University

Key Words: Fluctuation analysis, Luria-Delbruck distribution, Mutation rates

Fluctuation analysis is an established technique of estimating mutation rates in cultured cells. This paper presents a modified median estimator of mutation rate, which allows for large spread of the total culture population size N_t. Simulation results show a good accuracy and robustness of the modified median estimator compared with the traditional median estimator and the maximum likelihood estimator. The proposed estimator is applied to 116 experimental datasets collected for a study of chromosome loss and recombination in S. cerevisiae diploid yeast strains which vary in their genetic stability. An interesting empirical inverse relationship between the population size N_t and the mutation rate, which may cause inflation of variability of estimates, is discussed.

Robust Quantile Regression with Structured Design Space

❖ Ying Wei, Columbia University, 722 West 168th St Rm 629, New York, NY 10032, *yw2148@columbia.edu*; MatÌas Salibián-Barrera, The University of British Columbia

Key Words: quantile regression, robust, conditional screening

We propose a quantile estimate that is robust to two types of atypical observations in the design space: one is outlying covariates, known as high-leverage points; another type may not have manifested itself in any single component of X, but clearly deviated from the majority of the data. When the design space has certain structures other than ellipse, the second type outliers may be hidden by the existing algorithms assuming generic distribution. We propose a new data-adaptive method to quantify the deviations of a single observation in the design space, and incorporate them into the quantile regression estimation equations to achieve robust quantile estimates. The resulting quantile estimates have bounded influence, and are asymptotically consistent and normally distributed; the methodologies are applied to robustify conditional weight screening model based real Finnish growth data.

A Rank Regression Analysis of Hormonal and Cytokine Effects on Birth Outcome

❖ Alai Tan, The University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-1148, *altan@utmb.edu*; Roberta J. Ruiz, The University of Texas Medical Branch; Daniel H. Freeman, The University of Texas Medical Branch

Key Words: Rank regression, Linear regression, Log-transformation

Normality is one of the assumptions for traditional linear regression. However, this assumption is often violated in physiological data, particularly the cytokines and hormones. Despite this violation, traditional regression is regarded as a robust method when the sample size is relatively large. Another widely used method is regression on log-transformed variables. The present study compared the rank regression analysis with traditional regression and regression based on log-transformed variables in analyzing hormonal and cytokine effects on birth outcome. We found that rank regression is more powerful in detecting the effects of hormones and cytokines on birth outcome. We propose the use of rank regression analysis in analyzing hormone and cytokine data, especially when the traditional regression and log-transformed regression fail to detect the hypothesized associations.

Sequential Confidence Limits for the Ratio of Two Binomial Proportions

❖ Hokwon Cho, University of Nevada, Las Vegas, 4505 Maryland Parkway, Box 454020, Las Vegas, NV 89154-4020, *cho@unlv.nevada.edu*; Z. Govindarajulu, University of Kentucky

Key Words: Sequential confidence limits, ratio of two binomial proportions, maximum likelihood estimator

We present a sequential method for obtaining approximate confidence limits for the ratio of two independent binomial proportions based on the maximum likelihood estimator. Large-sample properties of the proposed sequential estimator are studied. Monte Carlo simulation is carried out in order to investigate its finite sample behavior. The proposed method is applied to a numerical example for illustration and its use.

An Index of Local Sensitivity to Nonignorability Under Pseudo-Likelihood

❖ Fang Zhu, University of Pittsburgh, Dept of Biostatistics Graduate School, 130 DeSoto Street, Pittsburgh, 15261, *faz10@pitt.edu*; Gong Tang, University of Pittsburgh

Key Words: missing data, sensitivity analysis, ISNI

Sensitivity analysis is recommended for regression analysis of data with nonresponse. Index of local sensitivity to nonignorability (ISNI) (Troxel, et al., 2004), was devised to detect the sensitivity of a parameter estimate to the ignorability assumption. However, ISNI requires specifying a parametric model for the missing-data mechanism. Here a new sensitivity index is proposed under a pseudo-likelihood method (Tang et al., 2003) that does not require specification of the mechanism. As an example, for bivariate data (x,y) when the mechanism is assumed to be an arbitrary function of x+lambda*y, this new index is defined as the 1st derivative of a parameter estimate as a function of lambda at lambda=0. For bivariate normal data a closed form of this new index was derived. Simulation studies have shown that this new index has comparable performance with the ISNI.

Fiducial Generalized Confidence Interval for Median Lethal Dose (LD50)

❖ Lidong E, Colorado State University, 120 Marble Street 302, Statistics Department, Broomfield, CO 80020, *e@stat.colostate.edu*; Jan Hannig, Colorado State University; Hariharan Iyer, Colorado State University

Key Words: LD50, logistic model, Fiducial Generalized Confidence Interval, Gibbs sampling

In this paper, we focus on the construction of confidence intervals for the median lethal dose LD50, which is a common measure of the acute toxicity of a compound in a species. We propose a new method for constructing confidence intervals of LD50 for a logistic-response curve. This method is based on an extension of R. A. Fisher's fiducial argument developed by Hannig (2006). The Gibbs sampling approach is used to empirically estimate the percentiles of the fiducial distribution for LD50. The resulting intervals are compared with three other competing confidence interval procedures - Delta method, Fieller's method and Likelihood Ratio method. Our simulation results show that fiducial intervals have satisfactory performance and are more stable than other confidence intervals in terms of coverage probability. The fiducial distributions also appear to give unbiased point estimates of LD50.

Multinomial Logistic Regression Models with a Stochastic Design Variable

❖ Evrim Oral, Middle East Technical University, Middle East Technical University, Department of Statistics, Ankara, 06531 Turkey, *eoral@metu.edu.tr*

Key Words: Multinomial Logistic Regression, Generalized Linear Models, Robust Inference, Modified Maximum Likelihood Method

Multinomial Logistic Regression has wide applications in many fields. Most of the literature on this topic depends on classical methods which assumes that the design variable X is non-stochastic in nature. In recent years, however, it has been recognized that non-stochasticity of the design variable is too restrictive and nonrealistic: see, Oral (2006). By the use of the modified maximum likelihood methodology, the model has been opened up to stochastic design variable. The aim of this study is to consider the estimation and hypothesis-testing problems in multinomial logistic regression models with a stochastic variable X. MML method is used in all of the theoretical derivations. MML estimators are known to be highly efficient and robust. Mathematical properties of these estimators and test statistics will be analyzed through simulations and real life examples will be carried out.

334 Adaptive Designs in Drug Development ● ✪

Biopharmaceutical Section, Biometrics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Findings from Simulation Studies on Treatment (Trt) Switching in Clinical Trial (CT)

❖ Dachuang Cao, Eli Lilly and Company, 1209 South College Street, APT 2314, Charlotte, NC 28203, *caoda@lilly.com*; Hongmei Jiang, Northwestern University,

Key Words: treatment switching, marginal structural model, generalized estimating equation (GEE), repeated measures analysis

Some CTs had the core study period (CSP) and extension period (EP), with subjects randomized to be treated or untreated in CSP, and untreated subjects allowed receiving trt in EP. To examine trt effect at the end of EP, some planned analyses compared endpoint between treated group (TG) and untreated group (UG) by defining TG as subjects who ever received trt. Also, baseline was defined differently: for a subject without trt switching, baseline was the start of CSP; otherwise, the start of EP. Simulation was performed to compare such planned analysis and other methods: repeated measures analysis (RM), GEE, and marginal structural model (MSM) with stabilized weight (SW) and nonstabilized weight (NSW). The results showed that MSM with SW and GEE were most reliable, that MSM with NSW and planned analysis often had large variation, and that it was difficult to interpret the results of RM.

Adaptive Design with Continuous Responses and Binary Utility

❖ Yuehui Wu, GlaxoSmithKline, 1250 S Collegeville Rd, PO Box 5089, Collegeville, PA 19426, *yuehui.2.wu@gsk.com*; Valeri V. Fedorov, GlaxoSmithKline

Key Words: multiple endpoints, utility function, optimal design, dichotomization

In clinical trials, correlated multiple endpoints problems are commonly seen. Sometimes, the observed responses are continuous but a regulatory agency will approve the drug only if the probability that the efficacy measure exceeds a predefined threshold while the toxicity does not exceed another given threshold. Thus the probability of interest (utility) is based on dichotomized responses. We emphasized the usefulness of differentiation between response and utility functions and developed tools to build fully adaptive and composite designs for the corresponding models. We use locally optimal designs as benchmarks. Two criteria: D-optimality and L-optimality (variance of estimated location) are considered. It is also shown that the practice of reporting dichotomized responses leads to a substantial loss in the precision of estimated parameter.

Midcourse Selection of the Primary Endpoint in Clinical Trials

❖ Zhilong Yuan, Johnson & Johnson PRD, 920 Route 202 South, PO Box 300, Raritan, NJ 08860, *zyuan1@prdus.jnj.com*; Yang Song, Johnson & Johnson PRD; Xiaolong Luo, Johnson & Johnson PRD; George Chi, Johnson & Johnson PRD

Key Words: Endpoint Selection, Adaptive Design, Clinical Trial, Type I Error Rate

Mid-course selection of endpoints is often encountered in clinical trials. Several group sequential and adaptive methods have been developed to address this issue but few has explicitly incorporated the selection rules, hence the loss of power. We formulate endpoint selection in the closed testing framework, and characterize the corresponding type I errors and powers. We then propose flexible alpha allocation and apply it to the global test to optimize the power, which is subsequently illustrated in a cancer clinical trial. A generalization to conditional error functions is also discussed which allows the incorporation of other flexible design features such as sample size re-estimation.

On Treatment Selection in Accelerated Drug Development

❖ Ying Wan, Temple University, Speakman Hall 006-00, 1810 North 13th Street, Philadelphia, PA 19122, *ying_wan@merck.com*; Keaven M. Anderson, Merck & Co., Inc.; Cong Chen, Merck & Co., Inc.; Boris Iglewicz, Temple University

Key Words: accelerated test designs, clinical trial, interim analysis, two-stage design

We present a method for the design of a clinical trial to combine phase 2 and 3 of clinical development. The method provides a two-stage study design allowing a fixed or data dependent number of treatment arms to be selected for stage 2. An example is given to illustrate the method and to examine advantages of the two-stage flexible design. We study some key questions in finding a good strategy for detecting the best treatment(s) and producing a significant result in consideration of limited resources for drug development. Given a prespecified fixed sample size, practical recommendations are made on: whether to select one arm or multiple arms, whether to expand the number of experimental arms in a trial, and what is the appropriate timing of interim analysis.

Adaptive Treatment Selection in Seamless Phase II/III Designs

❖ Xiaoling Wu, Bristol-Myers Squibb Company, 5 Research Parkway, Wallingford, CT 06492, *xiaoling.wu@bms.com*

Key Words: seamless adaptive design, dose selection, multiple testing

An adaptive seamless design combines a learn phase and a confirm phase together to reject or accept the global hypothesis. Adaptive treatment/dose selection is an important applications of seamless designs as poor dose selection and failure to collect information from early phase studies can cause negative result of a phase III clinical trial. Our work applies several closed multiple testing procedures to control the type I error rate in the setting of an adaptive two-stage design in which our primary interest is to compare multiple treatments with a control. We first review the general procedure to integrate Bonferroni-Holm type and Sims's test type multiple testing into a two stage adaptive dose selection design. Secondly, we use simulated examples to demonstrate advantages and disadvantages of such flexible design over traditional design (splitting learn phase and confirm phase).

Adaptive Design and Decision Analysis: Companions for Transforming Clinical Development

❖ Andreas Sashegyi, Eli Lilly and Company, 450 E Ohio Street, Unit 102, Indianapolis, IN 46204, *aisasheg@lilly.com*

Key Words: Adaptive Design, Decision Analysis, Clinical Trials

As the traditional drug development paradigm becomes increasingly less sustainable given today's challenges and expectations for the future, regulatory agencies and pharmaceutical developers are taking steps in support of modernizing and increasing the efficiency of drug development. Adaptive clinical trial designs have received much recent attention, holding considerable promise as a means of improving efficiency. In parallel, structured decision analysis as a discipline in its own right has the potential for more widespread application in clinical development than is currently witnessed. We highlight the natural connection between principles underlying both adaptive design and decision analysis, and offer practical examples of the synergistic application of these two methodological frameworks, emphasizing that the latter may be seen as an enabling tool for the former.

Unweighted versus Weighted Adaptive Design and Its Application in COPD Clinical Trials

❖Peter Zhang, Otsuka Pharmaceutical Development & Commercialization, Inc., 2440 Research Boulevard, Rockville, MD 20850, *peterz@otsuka.com*

Key Words: adaptive design, sample size, interim analysis, pivotal trials, COPD, Unweighted statistics

There are many so-called "adaptive design" for early phase proof of concept or dose selection studies. We focus on adaptive design for late phase pivotal trials seeking regulatory approval by FDA or EMEA. Although bio-tech and pharmaceutical companies are very good at to produce "successful" phase 2 trials the failure rate of phase 3 pivotal trials is very high and getting lower. We shows that adaptive design/re-estimating sample size at the interim analysis can short drug development cost and time by reducing the failure rate of Phase 3 trials when the drug is promising. Weighted statistics and Unweighted methods are compared with examples and simulations. An real example of adaptive/design in phase 2/3 pivotal studies in COPD will be provided.

335 Design Issues for Clinical Trials ●

ENAR, Biometrics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Correcting for Misrandomization in a Peer-Teaching Intervention Trial

❖Dirk Moore, University of Medicine and Dentistry of New Jersey, School of Public Health, 683 Hoes Lane West, PO Box 9, Piscataway, NJ 08854, *mooredf@umdnj.edu*; Weichung J. Shih, University of Medicine and Dentistry of New Jersey; Thomas Kearney, University of Medicine and Dentistry of New Jersey

Key Words: intervention trial, mis-randomization

A common design for an intervention trial is to randomly assign subjects into two groups. But in some cases, perfect randomization is not practicable, since at times the assigned intervention may not be available. For example, in a recently completed mammography peer teaching study, the objective was to compare the effectiveness of a nurse practitioner to that of a peer of the subject in explaining the importance of regular mammograms. The outcome was whether or not the subject returned for a mammogram within one year of the intervention. However, for logistical reasons, the randomly assigned intervention (nurse or peer) was not always available, and in such a case, the patient was re-assigned to the alternative intervention. In this talk we discuss using covariates, propensity scores, and instrument variables as potential tools for adjusting for the mis-randomization.

Evaluation of Surrogate Biomarker in a Randomized Clinical Trial

❖Long-Long Gao, GlaxoSmithKline, 485 Virginia Ave, Paoli, PA 19301, *long-long.2.gao@gsk.com*; Marshall Joffe, University of Pennsylvania; Xiwu Lin, GlaxoSmithKline

Key Words: Instrumental variables, randomized trial, surrogate biomarker

The statistical challenge of evaluating potential surrogates in a randomized trial is how to estimate the true relationship between biomarker and clinical endpoint since both biomarker and clinical endpoint are outcome variables. In this paper, we use direct and indirect effect concept (Pearl, 2002) to formalize the question, and propose a two-stage estimation approach to generalize the traditional instrumental variables method (Bowden, 1984). A simulation shows that our approach provides more consistent estimation than the proportion of treatment effect explained approach (Freedman, 1992). Our model assumptions are weaker than the sequential ignorability required by Freedman's approach. We also show the connection of our approach to the meta-analysis approach (Daniels, 1997) and the principal stratification method (Frangakis, 2002).

Proportion of Treatment Effect Explained by a Continuous Surrogate Marker

❖Jie Huang, Novartis Pharmaceuticals, 170 Park Avenue, Bldg 104 Room 2J20, Florham Park, NJ 07936-1080, *jie.huang@novartis.com*; Bin Huang, Cincinnati Children's Hospital Medical Center

Key Words: Surrogate, Prentice Criteria, Freedman's PTE, Wang and Taylor's F(F`), Probit Regression, Clinical Trial

Using surrogate endpoint in clinical trial is desirable for drug development because the trial will be cost-effective. Validating surrogate endpoint for clinical endpoint is critical to warrant such use. Since Prentice's definition (1989), different proposals were made to define proportion of treatment effect explained (PTE) by a surrogate. This paper extended Wang and Taylor's (2002) work for the case of continuous surrogate marker and binary outcome/treatment, which is one of the most common settings in clinical trial. We provided an easy and accurate solution to measure the proportion of treatment effect explained by a continuous surrogate endpoint. Further, we provided a simple graphical presentation of the surrogate effect. Both simulation results and actual trial example were produced to compare this alternative measure with Freedman's (1992) PTE measure.

Inference of Noncentrality Parameter of a Truncated Chi-Squared Distribution, with Application to Flexible Two-Stage Designs

❖Qizhai Li, National Institutes of Health, Academy of Mathematics and Systems Science, Biostatistics Branch, 6120 Executive Blvd, Executive Plaza South, Room 8047, DCEG, NCI, Rockville, MD 20852-7244, *liq@mail.nih.gov*; Kai Yu, National Institutes of Health

Key Words: Flexible design, non-centrality parameter, double bootstrap, non-central chi-squared distribution

We study the inference property of noncentrality parameter estimation based on samples from a truncated chi-squared distribution. This work is motivated by the flexible two-stage design using the log likelihood ratio test statistic, where the second-stage sample size is decided by the estimated noncentrality parameter of a truncated chi-squared distribution. We focus on the moment estimator due to its superior performance compared with other types of estimators, such as the MLE, and present theoretical results regarding to its existence and uniqueness. We propose a double bootstrap procedure for confidence interval constructions. Simulation studies are conducted to assess the accuracy of the moment estimation and evaluate the coverage property of the proposed confidence interval. Finally we show the application of the proposed method in the flexible two-stage design.

An Internal Pilot Study with Interim Analysis for Gaussian Linear Models

❖John Kairalla, The University of North Carolina at Chapel Hill, Dept. of Biostatistics, 123 Chaucer Court, Carrboro, NC 27510, *johnk@unc.edu*; Keith E. Muller, University of Florida; Christopher S. Coffey, The University of Alabama at Birmingham

Key Words: internal, pilot, adaptive, analysis, power, design

Combining early stopping rules with internal pilot methods would increase study flexibility, scope, and efficiency for general linear models. I first introduce the model and describe the procedure for a two-stage study design for Gaussian linear models with sample size recalculation based on an updated error variance estimate. Importantly, the design incorporates possible stopping for futility and efficacy at the first stage. Derivations of exact distributions in computable forms allow accurate calculations of power, type I error rate, and expected sample size. The theoretical results jointly cover first and second stage test statistics for both single d.f. hypotheses such as t-tests and more complex hypotheses including multiple group comparisons. The results can be used during planning to create an efficient study with early stopping rules and predictable power properties.

Estimation of Correlated Response Rates in Clinical Trials

❖Myron Chang, University of Florida, 1465 SW 88th St, Gainesville, FL 32607, *mchang@cog.ufl.edu*

Key Words: Self-consistent estimator, Asymptotic relative efficiency, Correlated responses, Information matrix

Clinical trials sometimes contain two or more binary response endpoints, such as hematological response, cytogenetic response, and molecular response. Usually, all response rates should be estimated and reported. In a clinical trial setting, these responses are often closely correlated and some patients are not evaluable for one or more responses due to uncontrollable factors, such as insufficient cell quantity in samples leading to indetermination of one or more responses. We propose a self-consistent estimator under the assumption that the response rates are the same for evaluable patients and for inevaluable patients. The proposed estimator is more efficient than the standard estimator in terms of asymptotic relative efficiency. Simulation results and a real example will be provided.

Test Size for Two-Stage and One-Stage Analyses of Cluster Samples of Unequal Size

❖Jacqueline Johnson, The University of North Carolina at Chapel Hill, 700 Bolinwood Drive Apt 24G, Chapel Hill, NC 27514, *jjohnson@bios.unc.edu*; Diane Catellier, The University of North Carolina at Chapel Hill; Keith E. Muller, University of Florida

Key Words: clustered data, group randomized trials, gaussian linear models

Important public health research often requires the use of community based studies due to logistical, ethical and cost constraints. Such designs require special methods of analysis. Gaussian clustered data are often analyzed with either a two-stage analysis of cluster means or a mixed effects linear model on individual level data. For data with a large number of clusters and large number of observations within each cluster, both techniques provide unbiased hypothesis tests. In small samples with unbalanced data, however, even moderate imbalance in cluster size across treatment groups can lead to test size bias. We present the results of a simulation study to describe bias in test size in the two-stage analysis of cluster means and one-stage mixed model for unbalanced clustered data.

336 MLE and Related Topics ✪

IMS, Section on Nonparametric Statistics

Tuesday, July 31, 2:00 pm–3:50 pm

Testing Homogeneity Variances for Non-Normal Models Using Entropy

❖Solomon W. Harrar, University of Montana, 32 Campus Drive, Department of Mathematical Science, Missoula, MT 59812, *harrar@mso.umt.edu*; Arjun K. Gupta, Bowling Green State University; Leandro Pardo, Complutense University of Madrid

Key Words: Sample Variance, Cumulants, Edgeworth Expansion, Characteristic Function, Robustness, Entropy

This talk is concerned with comparison of variability among populations which are not necessarily normally distributed. Our measure of variability is based on an information theoretic criteria known as Shannon's Entropy. The main problem with the standard methods is that they are sensitive to the distributional assumptions. Based on Shannon's Entropy, we derive a test for equality of variability assuming a general location scale family for the populations. Asymptotic expansion is obtained for the null distribution of the test statistic. The technique develops a multivariate Edgeworth expansion of the characteristic function and a formal inversion of the expanded characteristic function. Our test is shown to perform well compared to the likelihood ratio criterion when the populations have normal as well as non-normal distributions.

Bootstrapping the Grenander Estimator

❖Bodhisattva Sen, University of Michigan, 1085 South University, 439 West Hall, Ann Arbor, MI 48109, *bodhi@umich.edu*

Key Words: Consistency of bootstrap, Cube-root asymptotics, Empirical distribution function, Grenander estimator, Least concave majorant, Smoothed bootstrap

We consider the problem of bootstrapping the Grenander estimator, the nonparametric maximum likelihood estimator of f, a non-increasing density on positive real line. The Grenander estimator converges weakly to a non-normal limit involving nuisance parameters at $n^{1/3}$-rate. The non-standard rate of convergence makes the usual bootstrap procedures a suspect in this situation. In this paper we explore different bootstrap methods and investigate the consistency of the procedures. We show that the bootstrap statistic based on i.i.d. samples from the empirical distribution function, does not have any weak limit, conditional on the data, in probability. A similar phenomenon holds while bootstrapping from the least concave majorant of the empirical distribution. A suitable version of smoothed bootstrap achieves consistency, as does the m out of n bootstrap.

Estimation of Tail Probability via the Maximum Lq-Likelihood Method

❖Davide Ferrari, The University of Minnesota, 2632 35th Ave S, School of Statistics, Minneapolis, MN 55406, *dferrari@stat.umn.edu*

Key Words: Maximum Likelihood estimation, entropy, tail probability, rare events

Estimation of tail probability is of interest in various applications. Given a parametric model, a natural approach is maximum likelihood estimation. Although the resulting estimator is asymptotically efficient, the large sample property is often not trustworthy for estimating small tail probabilities. We introduce a new estimator for the parameters, the Maximum Lq-Likelihood Estimator (MLqE), based on Havrda and Charv·t entropy function, and apply it for estimating tail probabilities. Its behavior is characterized by the degree of distortion, q, applied to the assumed model; when q is close to 1 the new estimator approaches the usual MLE. We derive asymptotic properties of the MLqE and assess its efficiency, showing that it successfully trades bias for variance when the amount of information available is not large relative to the size of the tail probability to be estimated.

Minimax-Bias L-Estimators of Scale

❖ Daniela Szatmari-Voicu, University of Texas at El Paso, 6022 Caprock Ct Ap 102, El Paso, TX 79912, *dsvoicu@utep.edu*

Key Words: L-estimators, scale parameter, minimax-bias, robust estimation

We derive the best nonsymmetrized and symmetrized L-estimators of scale with respect to their asymptotic bias, under e-contamination neighborhood centered at a known error distribution Fo which is symmetric and unimodal, by first performing a study on the maximum asymptotic bias curves of the nonsymmetrized and symmetrized interquantile ranges from which the L-estimators considered arise. For Fo symmetric and unimodal and making use of the generalized method of moment spaces, the solutions to the minimax-bias problems are then shown to be convex combinations of at most two interquantile ranges.

Asymptotic Properties of MLEs for Extended Generalized Log-Logistic Families of Lifetime Distributions

❖ James U. Gleaton, University of North Florida, 4567 St Johns Bluff Road South, Jacksonville, FL 32224, *jgleaton@unf.edu*; M. Mahbubur Rahman, University of North Florida

Key Words: extended generalized log-logistic transformation, maximum likelihood estimator, asymptotic properties

An EGLL family of lifetime distributions is one in which any pair of distributions are related through an EGLL transformation, which is a composition of a proportional odds transformation and a generalized log-logistic transformation. It may be that one member of the family has a simple form, e.g., two-parameter Weibull, and may be considered the generating member. Any other member is called a composite distribution, having two additional parameters, due to the transformation. It is shown that if the generating distribution has certain regularity properties, then the MLEs for the parameters of a composite distribution have the properties of consistency and asymptotic normality and efficiency, for a certain range of values of the transformation parameters.

Profile Likelihood-Based Confidence Intervals in Repeated Categorical Data

❖ Esa I. Uusipaikka, University of Turku, Department of Statistics, Assistentinkatu 7, Turku, 20014 Finland, *esa.uusipaikka@utu.fi*

Key Words: Profile likelihood, Confidence interval, Repeated measurements, Categorical data

Profile likelihood-based confidence intervals for components or more generally real-valued functions of parameter vector are approximate confidence intervals based on likelihood ratio statistic and its asymptotic chi square-distribution. The calculation of these intervals is possible if explicit expressions of the log-likelihood and score functions are available. Joint regression and association modeling of repeated categorical responses using dependence ratios provides explicit expressions for log-likelihood and score functions and thus the above mentioned way of constructing confidence intervals is available. The method is illustrated by constructing profile likelihood based confidence intervals for several parameter functions of interest in longitudinal dataset reporting the presence and absence of wheeze at ages 7, 8, 9, and 10 among 537 Steubenville children.

Maximum Likelihood Estimation of the Dispersion Parameter of Negative Multinomial Distribution

❖ Soumi Lahiri, GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, PA 19426, *soumi_lahiri@hotmail.com*

Key Words: negative Multinomial Distribution, Maximum Likelihood Estimator, EM Algorithm

The negative multinomial sampling plans are widely used in the biomedical field, especially modeling rare events. A major drawback to the use of this distribution is the difficulty in estimating the shape parameter k. The maximum likelihood estimator (MLE) of k cannot be obtained by directly solving the likelihood equation for the negative multinomial distribution. This paper describes a maximum likelihood method using the EM algorithm to estimate k. The efficiency of this method has been illustrated with simulation results.

337 Bayesian Approaches to Sample Size Determination and Experimental Design ●

Section on Bayesian Statistical Science

Tuesday, July 31, 2:00 pm–3:50 pm

Survey of Bayesian Sample Size Calculation

❖ Susan Alber, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 237, Houston, TX 77030, *salber@mdanderson.org*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: Bayesian clinical trials, prior specification, operating characteristics

Sample size calculation is an essential step in designing clinical trials. Under the frequentist's framework, methods for computing sample size are well established and there are many commercial and free software packages for implementing these methods. However under the Bayesian framework there is a lacking in both the theoretical development and software availability for sample size calculation. We survey the literature and compare published methods for Bayesian sample size calculation. We identify the key components and propose a general approach for computing sample size under the Bayesian framework. Methods will be discussed to achieve two goals: 1. a predetermined accuracy for estimation and 2. desirable operating characteristics for hypothesis testing.

Bayesian Power and Sample Size Calculations for Multicenter Clinical Trials with Binary Responses

❖ Emine O. Bayman, The University of Iowa/Uludag University, 4155 WL, Dept of Biostatistics, Iowa City, IA 52242, *emine-unlu@uiowa.edu*; Kathryn Chaloner, The University of Iowa; Mary K. Cowles, The University of Iowa

Key Words: Bayesian power, Bayesian sample size, Multi-center clinical trials, Binary response

An important issue is to decide the total sample size and sample size for each center in a multi-center clinical trial. A Bayesian hierarchical model for binary responses representing success or failure is assumed, with exchangeable random center effects. A concept of Bayesian power is defined as the probability of reaching a specific conclusion for a fixed set of parameter values. An algorithm is described for determining the sample size to give appropriate Bayesian power. It is assumed that there are 2 treatments and N centers. Each center has a pre-specified relative ability to enroll. In addition, the impact of balance and imbalance between sample sizes in each center on power is examined. The algorithm is implemented using WinBUGS and R, and the R2WinBUGS interface.

Bayesian Sample Size Determination for the Misclassified Binary Response Variable with One Error-Free Covariate

❖ Dunlei Cheng, Baylor University, 1501 S. 9th Street, Apt. 402, Waco, TX 76706, *dunlei_cheng@baylor.edu*; James Stamey, Baylor University

Key Words: sample size calculation, Bayesian inference, sampling and fitting priors, Misclassification, logistic regression

The paper studies Bayesian simple size determination for a misclassified binary response variable when one error-free covariate is included. The performance criterion employed in this study is Bayesian power. The parameter of interest is the covariate direction in the logistic regression model. Unlike many occasions where priors for simulating data and posterior distribution are identical, this study uses more informative sampling priors in the data collection whereas vague fitting priors in the posterior simulation. Simulations demonstrate that the choice of sample priors have a great impact on sample size calculation. Moreover, multiple diagnostic tests result in a smaller sample size.

Bayesian Sample Size Calculation for Counted Data with Excess Zeroes

❖ Chunyao Feng, Amgen Inc., One Amgen Center Drive, MS 24 2 A, Thousand Oaks, CA 91320, *afeng@amgen.com*; John W. Seaman, Jr., Baylor University; James Stamey, Baylor University

Key Words: Bayesian, Sample Size, Negative Binomial, Zero-inflated Poisson, Generalized Poisson, Model

Hospitalization is the largest portion of the total cost of patients to the health services. Interest lies in comparing the rate of hospitalization for two groups of patients. We consider a Bayesian method to determine the required sample size to test equality of these rates with a pre-specified power. Generalized Poisson (GP), Negative Binomial (NB) models, Zero-inflated Poisson (ZIP) and Zero-inflated Negative Binomial (ZINB) models are considered as the candidates to model the rate of hospitalization. In addition, we compare the performance of these models via Bayesian model comparison techniques. Finally, a sensitivity analysis is performed to test the robustness of this methodology.

Bayesian Up and Down: Combining the Best of Both Worlds

❖ Assaf Oron, University of Washington, 4200 Mary Gate Mem Dr NE, Apt P217, Seattle, WA 98105, *assaf@u.washington.edu*; Peter D. Hoff, University of Washington

Key Words: Quantile Estimation, Up-and-Down, CRM, Bayesian Inference, Phase I Clinical Trials, Sensory Studies

Percentile-finding sequential designs aim to concentrate treatment allocation around the target percentile Q_p, as quickly and efficiently as possible. The 'Up-and-Down' (U&D) algorithm utilizes the Markov property, to converge with a geometric rate to a stationary treatment distribution peaked near Q_p. The Bayesian CRM approach, developed in the context of Phase I clinical trials, utilizes a parametric model to achieve better asymptotic allocation concentration than U&D. However, CRM's rate of convergence is $\sqrt{n}$ at best. We propose to combine the two approaches within a Bayesian framework. The U&D transition rule is used as a default, to be overridden only when the CRM transition promises improvement with acceptable posterior credibility. Simulation results demonstrate the potential of this approach for Phase I and other percentile-finding applications.

How Many Images Are Enough? Bayesian and Non-Bayesian Approaches to Sample Size Determination

❖ Junheng Ma, Case Western Reserve University, Apt D 106, 5506 Fifth Ave, Pittsburgh, PA 15232, *jxm216@case.edu*; Jiayang Sun, Case Western Reserve University; Joe Sedransk, Case Western Reserve University

Key Words: posterior predictive p-value, power calculation, sample size, images, Multinomial distribution

In developing new imaging diagnostic tools, it is often of interest to compare the quality of the resulting images from the competing tools. Each image is assessed by one or several evaluators who score the images from 1 to 5. An important scientific question is how many images are needed to detect differences in the quality of these images when using these diagnostic tools. Specifically, one is interested in finding the sample sizes needed to detect a difference between the parameters from two multinomial populations. In this paper we develop three Bayesian approaches based on the ideas of posterior predictive p-values and two non-Bayesian methods using a chi-square test and a simultaneous confidence interval. Evaluations and comparisons are made using asymptotics, simulated data and a real data application.

Kullback-Leibler Divergence for Design and Model Selection

❖ Chenpin Wang, University of Texas Health Science Center San Antonio, 13507 Charter Bend Dr, San Antonio, TX 78231, *wangc3@uthscsa.edu*

Key Words: Kullback-Leibler Divergence, Fisher information, optimal design, model comparison

We consider two Kullback-Leibler distances (KLDs) suitable for comparing model fit or study designs in the Bayesian framework: (a) the posterior mean of a weighted KLD between the predictive distribution of a diagnostic statistic T_n under a model g with a correct specification of its first two moments and that under an assumed model f, and (b) the KLD between the posterior distributions of certain model parameters associated with T_n as described in (a). We prove the asymptotic equivalence between the KLDs in (a) and (b) under certain regularity conditions and that the estimation of the first moment of T_n is unbiased under the assumed f. Superiority of these two KLDs to both the Bayesian and frequentist D-optimal criteria is demonstrated with examples when the assumed f is subject to biased estimation of the second moment of T_n, but not its first moment.

338 Resampling Methods and Confidence Intervals ✪

Section on Nonparametric Statistics

Tuesday, July 31, 2:00 pm–3:50 pm

Bootstrap Confidence Intervals in Nonparametric Regression with Built-In Bias Correction

❖ McMurry Timothy, DePaul University, Mathematical Sciences, 2320 N Kenmore Ave, Chicago, IL 60613, *tmcmurry@depaul.edu*; Dimitris Politis, University of California, San Diego

Key Words: Nonparametric Regression, Bootstrap Confidece Intervals, Fourier Transform, Bias Reduction

The problem of estimating nonparametric regression with associated confidence intervals is addressed. It is shown that through appropriate choice of infinite order kernel, it is possible to construct bootstrap confidence intervals which do not require either explicit bias correction or suboptimal levels of smoothing at any stage of the estimation. In particular, it is demonstrated that in this setting, consistent estimates are obtained when both the pilot and final smoothings are estimated at the mean square error optimal bandwidth for estimating the regression. The effectiveness of the method is demonstrated through a small simulation study.

Tapered Empirical Likelihood Methods

❖Daniel Nordman, Iowa State University, 315 F Snedecor Hall, Department of Statistics, Ames, IA 50011, *dnordman@iastate.edu*

Key Words: block, periodogram, data taper

The talk addresses two different versions of empirical likelihood (EL) for time series, where both versions involve data tapers. The first formulation involves a ``time domain” EL based on tapered data blocks. Tapering helps to reduce the bias incurred in the studentization steps internal to blockwise EL. Simulation results indicate the data taper improves the coverage accuracy of EL confidence intervals for means. The second tapered version of EL is based on a tapered periodogram for inference on ``frequency domain” parameters, such as autocorrelations. This EL formulation involves no block choices, but the approach is valid for a smaller class of time processes and parameters. Simulation also indicates that tapering increases EL coverage accuracy in the frequency domain.

An Empirical Likelihood-Based Approach To Estimate Covariance Matrices with Structural Zeroes

❖Sanjay Chaudhuri, National University of Singapore, Dept of Statistics and Appl Prob, Faculty of Science 6 science drive 2, Singapore, 117546 Singapore, *sanjay@stat.nus.edu.sg*

Key Words: empirical likelihood, covariance matrices, covariance graphs, constrained maximisation

Zeroes in the covariance matrices of a multivariate random vector indicates linear independence between the corresponding components. Due to the presence of structural zeroes a maximum likelihood estimator cannot be expressed analytically and can only be conveniently iteratively computed, if the underlying distribution is Gaussian. Currently same or related algorithms are used to compute an estimator if the underlying distribution is non-Gaussian. In this talk we present an empirical likelihood based estimate of the covariance matrices, which preserves the structural zeroes. This estimator is unique, consistent and convenient to apply in a variety of situations. Moreover we shall show that, it is more efficient than the other estimators for non-Gaussian case and slightly less efficient than the mle for the Gaussian case. Some generalizations of the methodology will also be discussed.

Conditional Empirical Likelihood Inference in Presence of Nuisance Parameters

❖Mi-Ok Kim, Cincinnati Children’s Hospital Medical Center, MLC 5041 3333 Burnet Ave, Cincinnati, OH 45229-3039, *miok.kim@cchmc.org*

Key Words: empirical likelihood, nuisance parameter, censored data

Empirical likelihood (EL) is a nonparametric inference method with results that are in general similar to those about likelihood ratio tests and Wilk’s theorem in the parametric model. In this talk we formulate EL inference in the frame work of ordinary parametric likelihood using a parameterized sub-family of distributions and discuss its inference. In particular we discuss EL inference in presence of nuisance parameters. We propose an approximate conditional EL inference and extend this proposition to censored case.

Strong Approximations for Resample Quantile Processes and Application to ROC Methodology

❖Jiezhun Gu, North Carolina State University, 625 Chappell Dr, Raleigh, NC 27606, *sherrygu2001@yahoo.com*; Subhashis Ghoshal, North Carolina State University

Key Words: Bayesian bootstrap, Bootstrap, Empirical process, Kiefer process, ROC curve, Strong approximations

ROC curve is defined as true positive rate versus false positive rate obtained by varying a decision threshold criterion. Mathematically speaking, ROC curve is a composition of survival function of one population to the quantile function of another population. In this paper, we study strong approximation for the quantile processes of bootstrap and the Bayesian bootstrap resampling distributions, and use this result to study strong approximations for the empirical ROC estimator, the corresponding bootstrap, and the Bayesian versions in terms of two independent Kiefer processes. The results imply asymptotically accurate coverage probabilities for bootstrap and the Bayesian bootstrap confidence bands, and accurate frequentist coverage probabilities of bootstrap and the Bayesian bootstrap confidence intervals for the area under the curve functional of the ROC.

Observed Confidence Levels for Modes of Smooth Densities

❖Alan Polansky, Northern Illinois University, Division of Statistics, De Kalb, IL 60113, *polansky@math.niu.edu*

Key Words: bootstrap, kernel, functional, smoothing

The determination of the number of modes of an unknown smooth density is considered through the use of observed confidence levels. Observed confidence levels assign a level of confidence to hypotheses about the number of modes. Due to theoretical limitations, confidence levels can be assigned to only certain hypotheses about the number of modes. Methodology for computing the confidence levels is presented and is studied empirically.

Semiparametric Sequential D-Optimal Design

❖Joseph Warfield, Johns Hopkins University, 11100 Johns Hopkins Rd, Applied Physics Laboratory, Laurel, MD 20723, *joseph.warfield@jhuapl.edu*; Anindya Roy, University of Maryland, Baltimore County

Key Words: Constrained optimal design, Isotonic Smoothing spline, Link function, Sequential Design, Phase I Clinical Trial

Sequential D-optimal designs for binary regression models are commonly used in Phase I clinical trials and efficient estimation quantiles of the dose-response curve given by the link function. However, in most applications the link function is assumed to be known up to a location and a scale. We perform an empirical study to illustrate how severely the sequential design and the dose-response quantile estimation maybe affected when the link function is misspecified. We propose nonparametric estimation of the link function via isotonic smoothing spline estimator under monotonicity constraint and incorporate the estimator into the sequential allocation scheme. This makes the procedure more robust against model misspecification and at the same time maintains the objective of efficient estimation of the quantiles. The methodology is applied to data from a phase I clinical trial.

339 Financial Risk Analysis: Theory, Methods, and Applications

Section on Risk Analysis, Business and Economics Statistics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Extreme Value Theorem and Its Applications

❖Yan Li, The Pennsylvania State University, Statistics Department, 265 Blue Course Dr, Apt 8A, State College, PA 16803, *yul135@psu.edu*

Key Words: EVT, Operational risk, POT

Extreme risks are the concern of risk managers and investors. In order to estimate and hedge the risks better, Extreme Value Theory (EVT) has been fully developed in the past fifty years. The modeling framework based on the operational risk scenario will be established. The heavily skewed loss data histogram suggests a fat-tailed distribution such as the lognormal or Pareto distribution. Besides those standard parametric distributions, more advanced technique focusing on tail behavior, EVT, provides a better approach to fit the skewed data. Two principals of EVT, Block Maxima Models and Peaks-over-Thresholds Model, will be presented. The important asymptotic theories and the estimation procedures will be briefly introduced as well.

A New Estimator for the Extremal Index Based on Scaling and Resampling

❖Kam Hamidieh, University of Michigan, 528 S State, 152, Ann Arbor, MI 48104, *khamidie@umich.edu*; Stilian Stoev, University of Michigan; George Michailidis, The University of Michigan

Key Words: Extremal Index, Clusters, Heavy Tails, Extreme Value Theory

The extremal index is the key parameter quantifying the clustering of the extremes of a stationary sequence. In this talk, we propose a new estimator of the extremal index based on the scaling properties of the block maxima of heavy tailed data. We discuss some of its properties and demonstrate its competitive nature through an extensive simulation study. An application that explores the structure of a financial data set is also presented. This is joint work with Stilian Stoev and George Michailidis.

Tail Expansions for the Distribution of the Maxima of Random Walks with Negative Drift and Regularly Varying Increments

❖Chenhua Zhang, University of Georgia, Department of Statistics, Athens, GA 30602, *czhang@stat.uga.edu*; Philippe Barbe, CNRS, France; William P. McCormick, University of Georgia

Key Words: tail expansion, random walk, regularly variation, Wiener-Hopf factor, ruin probability

Let F be a distribution function with negative mean and regularly varying right tail. Under a mild smoothness condition we derive higher order asymptotic expansions for the tail distribution of the maxima of the random walk generated by F. The expansion is based on an expansion for the right Wiener-Hopf factor which we derive first. An application to ruin probability is developed.

A General Rule To Average a Set of Frequency Statistics

❖Joseph Kahan, FM Global, 1151 Boston Providence Turnpike, PO Box 9102, Norwood, MA 02062, *joseph.kahan@fmglobal.com*

Key Words: frequency, regression, heteroskedasticity

For the purpose of the presentation, a frequency statistic is defined by the number of losses divided by the total number of locations from which a loss could occur. Frequency statistics are estimated for a number of groups. In this presentation is proposed the best way of combining these groups to produce an average frequency statistic, where best is defined as being unbiased with a minimum variance. It is posited that a general regression model can form the structure through which three estimators are identified to be possible candidates for obtaining the average frequency of loss over all the loss groups. The best choice depends critically on the heteroskedasticity of the errors. The variances of these estimators depend critically on the correct specification of the model.

Bootstrapping the Expected Shortfall

❖Shuxia Sun, Wright State University, Deparment of Mathematics and Statistics, 3640 Col Glenn Hwy, Dayton, OH 45435, *shuxia.sun@wright.edu*

Key Words: expected shortfall, high quantile, risk measure, blocking bootstrap, nonparametric

The expected shortfall is a popular risk measure in financial risk management. It is defined as the expected loss on a portfolio of financial assets knowing that the loss is larger than a given high quantile. We derive the asymptotic properties of the blocking bootstrap estimators for the expected shortfall of a stationary process under strongly mixing conditions. Results from empirical examples and a small simulation study will also be presented to evaluate the performance of the proposed block bootstrap estimators.

Using Data Mining Tools and GIS (Geographic Information Systems) in Hot-Spot Detection

❖Wei-hong Wang, The College of New Jersey, Department of Mathematics and Statistics, Ewing, NJ 08628-4700, *wang@tcnj.edu*; Ben Liu, William Paterson University

Key Words: Data mining, Predictive modeling, Neural Network, GIS, Decision Tree, Logistic Regression

This study involves a specific data set which has 1.4 million cases, 14 predictors and one target variable with binary outcomes. The response surfaces and the classification accuracies of logistic regression and neural network models are relatively similar, yet the financial implications can be very profound when the models select the top 10% of cases and when the cost and profit are incorporated in the calculation. In addition, the Decision Tree and Geographic Information Systems (GIS) are used for case selection to enhance the profit to 96% of its full potential.

Estimation for Non-Negative, Levy-Driven Ornstein-Uhlenbeck Processes

❖Yu Yang, Colorado State University, 1500 W Plum St, Apt 6E, Fort Collins, CO 80521, *vicky_yangcsu@yahoo.com*; Peter J. Brockwell, Colorado State University; Richard A. Davis, Colorado State University

Key Words: continuous-time autoregression, Ornstein-Uhlenbeck process, Levy process, stochastic differential equation, sampled process, stochastic volatility model

The Ornstein-Uhlenbeck process (or stationary continuous-time autoregression of order 1 (i.e., CAR(1)) driven by nondecreasing Levy process has been used to model stochastic volatility of the log prices of financial assets (e.g., Barndorff-Nielsen and Shephard (2001)). In this talk, I will present a highly efficient method of estimation for the parameters of a CAR(1), taking advantage of the non-negativity of the driving process. I will also show how to reconstruct the background driving Levy process from a continuously observed realization of the process and use this result to estimate the increments of the Levy process itself when closely-spaced observations are available. Lastly, I derive the asymptotic distribution of the coefficient estimator for a gamma-driven CAR(1) and illustrate the performance of the procedure through a simulation study.

340 Clustering and Classification ●

Section on Statistical Computing, Section on Statistical Graphics, IMS

Tuesday, July 31, 2:00 pm–3:50 pm

Clustering the Short Stories of Edgar Allan Poe Using Galois Lattices

❖ Roger Bilisoly, Central Connecticut State University, 395 Brittany Farms Rd, Apt 437, New Britain, CT 06053, *bilisolyr@ccsu.edu*

Key Words: text mining, clustering, formal concepts, galois lattices

Edgar Allan Poe wrote almost seventy short stories in his lifetime, and literary critics have categorized these stories in many ways, (e.g., by genres such as horror, detective, or proto-science fiction). This talk discusses how a computer can group stories by using families of words related by a theme, e.g., words denoting colors. This approach combines two different techniques. First, we use term-document matrices, which were originally developed for document searches in the field of information retrieval. Second, we use formal concept theory, which defines concepts in a way that forms a Galois lattice. These lattices have both a well developed mathematical basis and have been used in applications beyond the computer sciences, (e.g., social networks in mathematical sociology). Finally, we will discuss how meaningful these groups of stories are to a human reader.

Smoothing the Dissimilarities Among Binary Data for Cluster Analysis

❖ David Hitchcock, University of South Carolina, Department of Statistics, 216 LeConte College, Columbia, SC 29208, *hitchcock@stat.sc.edu*; Zhimin Chen, University of South Carolina

Key Words: binary data, cluster analysis, smoothing, shrinkage, dissimilarity

Cluster analysis attempts to group data objects into homogeneous clusters on the basis of the pairwise dissimilarities among the objects. When the data contain noise, we might consider performing a smoothing operation, either on the data themselves or on the dissimilarities, before implementing the clustering algorithm. Possible benefits to such pre-smoothing are discussed in the context of binary data. We suggest a method for cluster analysis of binary data based on ``smoothed'' dissimilarities. The smoothing method presented borrows ideas from shrinkage estimation of cell probabilities. Some initial results are given, and some future avenues in this area are outlined. The method is illustrated with an example involving binary item response data.

Measuring the Similarity between Spatial Clusters

❖ Xinxin Bai, IBM China Research Laboratory, Building 19 Zhongguancun Software Park, Beijing, 100094 China, *baixx@cn.ibm.com*; Wen Jun Yin, IBM China Research Laboratory; Jin Dong, IBM China Research Laboratory

Key Words: Similarity Measure, Spatial Cluster, Feature Structure

To measure the similarity between two spatial clusters more reasonably, the special characteristics should be considered during spatial feature selection and feature weighting. One is that the features in one spatial cluster are often not independent of each other, therefore feature structure could provide valuable information. Another one, especially for the purpose of many real applications, is that some goal-related metrics are usually attached to spatial clusters and the similarity measures should align with those pre-defined goals as much as possible. In this paper, two spatial Feature Structure Similarity measures are firstly proposed. Then an optimization approach is employed to incorporate both feature structure and feature value characteristics with proper weights. Finally, the proposed similarity measures are tested in a real banking case for branch investment decision support.

Uniformity Testing in High Dimensions

❖ Adam Petrie, Rensselaer Polytechnic Institute, Troy, NY 12180, *petria@rpi.edu*

Key Words: Multivariate test of uniformity, High dimensional data, Minimum spanning tree, Hamiltonian Path, Hypercube, Clustering

An important yet not often studied problem in high-dimensional data analysis is determining whether data points are uniformly distributed over some compact domain, or rather possess some underlying structure instead (e.g., clumps, clusters, or other nonhomogeneities). We use the total length of the minimum spanning tree to confirm or reject uniformity in the unit hypercube. We also consider the number of runs and maximum run length of the sequence of segment lengths of a short Hamiltonian path (or "snake") constructed on the data to confirm or reject uniformity on an arbitrarily shaped compact domain in R^d. We find that both methods can detect clumps or clusters in otherwise uniformly distributed data with higher power than recent alternative methods proposed in the literature.

Presence versus Significance: How Significant Are Your Clusters?

❖ Rebecca Nugent, Carnegie Mellon University, Department of Statistics Baker Hall 132, Pittsburgh, PA 15217, *rnugent@stat.cmu.edu*; Werner Stuetzle, University of Washington

Key Words: clustering, confidence, bootstrap, level sets, generalized single linkage

The goal of clustering is to identify distinct groups in a dataset and assign a group label to each observation. To cast clustering as a statistical problem, we regard the data as a sample from an unknown population p(x). However, clustering methods rarely generate a one-to-one mapping of clusters to groups in the population. Groups may be partitioned into several clusters; spurious clusters may be falsely identified. A cluster's presence does not imply its significance. We introduce a bootstrap-based simultaneous confidence band used to estimate the hierarchical cluster structure of p(x) by the cluster tree of its level sets. Clustering with Confidence assigns a significance level to the cluster tree (and the individual clusters). Results for a graph-based estimation approach, generalized single linkage, will be shown.

Clustering Gene Expressions in the Presence of Scatter

❖ Ivan Ramler, Iowa State University, 124 Snedecor Hall, Ames, IA 50011-1219, *ramleri@iastate.edu*; Ranjan Maitra, Iowa State University

Key Words: Gene Expression, k-mean-directions, Bayes Information Criterion

A new methodology is proposed for clustering gene expression datasets in the presence of scattered observations. These are defined to be observations that are unlike any other, so traditional approaches that force them into groups can lead to all-around erroneous conclusions. Our suggested approach is an iterative scheme which proceeds by building the core for each cluster around the centers, identifies points outside as scatter and updates the method until convergence. In the absence of scatter, the algorithm reduces to a k-means algorithm designed for constrained directional data. We also provide methodology to initialize the algorithm as well as to estimate the number of clusters in the dataset. Results on several sets of test experiments show excellent performance. The methodology is applied to gene expression data on the diurnal starch cycle of Arabidposis L. Heynth.

A New Clustering Algorithm Based on Self-Updating Process

❖ Ting-Li Chen, Academia Sinica, Institute of Statistical Science, 128, Academia Rd. Sec. 2, Taipei 115, 115 Taiwan, *tlchen@stat.sinica.edu.tw*; Shang-Ying Shiu, Academia Sinica

Key Words: clustering, k-means, self-organizing maps, image segmentation

Many of the popular clustering methods, such as K-means and Self-Organizing Maps, require a set of initial values to begin the iterative process. In this talk we will present a simple and novel method that does not require such an initial set and can avoid the problem of local minima. The clustering strategy we propose is motivated by intuition on clustering. The algorithm stands from the viewpoint of subjects to be clustered and simulates the process of how they perform self-clustering. At the end of the process subjects belonged to the same cluster would converge to the same point, which represents the cluster location in a p-dimensional space. Our simulation study showed promising results compared to other clustering methods. An example on image segmentation will also be presented.

341 Adjustment Methods for Survey Nonresponse and Noncoverage ●

Section on Survey Research Methods, Social Statistics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Nonresponse Adjustment for Multipurpose Surveys

❖ Hee-Choon Shin, NORC at the University of Chicago, 2232 University Dr, Naperville, IL 60565, *shinh@uchicago.edu*

Key Words: Non-response Adjustment, Multi-purpose Survey, Weighting

The most common method of adjustment for unit non-response is weighting, where respondents and non-respondents are classified into adjustment cells based on some covariates known for all units in the sample, and a nonresponse adjustment factor is computed for all responding cases in each cell proportional to the inverse of the response rate in the cell. All the methods for non-response adjustment were developed for a single response or dependent variable. For example, the common predictive mean or response propensity stratification methods for constructing adjustment cells are based on a single outcome variable. We will describe a method for non-response adjustment for multi-purpose surveys in which there are hundreds of survey items.

Comparison of Methods to Adjust Further for Nonresponse Due to Missing Lab Data in NHANES

❖ Maya Sternberg, Centers for Disease Control and Prevention, 1600 Clifton Road NE, MS E63, Atlanta, GA 30333, *mrs7@cdc.gov*; Alula Hadgu, Centers for Disease Control and Prevention

Key Words: nonresponse, complex survey, NHANES, logistic regression, weighting

This paper compares the impact of additional nonresponse adjustments to the standard National Health and Nutrition Examination Survey (NHANES) weights using weighting class adjustments and logistic propensity models to estimate the prevalence of six distinct sexually transmitted diseases (STDs). Specimens to test for six STDs were collected on different subgroups of participants during different survey cycles, and using different specimen collection methods depending on the STD. We investigated the impact of missing lab test results on prevalence estimates of each of these STDs by comparing the results of weighting class adjustments, weighted and unweighted logistic propensity models, and further post-stratification and trimming adjustments. Based on these results we make recommendations about the impact of the weighting adjustments on descriptive inferences for STD prevalence.

Post-Katrina New Orleans Adjustment to Control Totals for the National Immunization Survey

❖ Steven Pedlow, NORC at the University of Chicago, 55 E Monroe, Chicago, IL 60603, *pedlow-steven@norc.uchicago.edu*; Zhen Zhao, Centers for Disease Control and Prevention

Key Words: Migration, Hurricane Katrina, Post-Stratification Totals

The National Immunization Survey (NIS), sponsored by the Centers for Disease Control and Prevention, is an ongoing telephone survey that measures vaccination coverage for children 19–35 months old. Interviewing resumed after Hurricane Katrina, but the population in Orleans Parish and the surrounding areas remains reduced. NIS weighting control totals are built from birth records that predate the Hurricane. As part of post-stratification, we created a special Orleans Parish out-migration step. This paper describes our use of special population estimates by the Census Bureau and the State of Louisiana's 2006 Louisiana Health and Population Survey, as well as data on Katrina aid applicants from the Federal Emergency Management Agency. With these data, we were able to realistically move the Orleans Parish population affected by Hurricane Katrina to their resettlement areas.

Diagnostic Process to Assess the Effects of Weighting BRFSS Data

Key Words: BRFSS, Weighting-Class Variables, Nonresponse Bias, Nonsampling Error, Standard Error

This research focuses on developing a diagnostic process that assesses the accuracy, robustness, and precision of the BRFSS prevalence estimates. This diagnostic process has three distinct sequential assessments: accuracy, robustness, precision. This diagnostic process serves to enhance the present BRFSS data quality monitoring processes. The 2004 NC BRFSS yielded accurate estimates of the proportion of: adult males and females, adults ages 18-34, 35-54, and 55+; less accurate estimates of the proportion of Hispanic adults, married adults, and adult high school graduates were also observed. Overall, robust and equally precise prevalence estimates were observed for six of the eight selected public health outcomes of interest. These robust and equally precise prevalence estimates provided evidence of negligible effects of errors due to disproportionate sampling rates, on response bias.

Minimizing Conditional Local Bias for Health Estimates from the Behavioral Risk Factor Surveillance System for U.S. Counties Contiguous to the United States-Mexico Border

❖ Joe Fred Gonzalez, Jr., Centers for Disease Control and Prevention, National Center for Health Statistics, 3311 Toledo Rd Room 3121, Hyattsville, MD 20782, *jgonzalez@cdc.gov*; Machell Town, Centers for Disease Control and Prevention; Jay H. Kim, National Center for Health Statistics

Key Words: unbiased estimation, poststratification

The Behavioral Risk Factor Surveillance System (BRFSS) is a State telephone based survey of the adult civilian non-institutionalized population residing in the United States. Consequently, the BRFSS final weights that are currently available in the data files are designed to produce unbiased estimates of socio-demographic and health characteristics at the State level.

In addition to State level BRFSS estimates, there is interest in the health status of adults residing in the 25 U.S. counties contiguous to the United States-Mexico Border Region. The purpose of this paper is to apply an alternative approach for poststratification by minimizing local (cell) bias of BRFSS estimates of socio-demographic and health characteristics of adults residing in the 25 counties contiguous to the U.S.-Mexico border.

Evaluation of the Current Weighting Methodology for the Behavioral Risk Factor Surveillance System (BRFSS) and Alternatives for Improvement

❖Mansour Fahimi, RTI International, 6110 Executive Boulevard Suite 902, Rockville, MD 20852-3907, *fahimi@rti.org*; Darryl Creel, RTI International; Paul S. Levy, RTI International; Michael Link, Centers for Disease Control and Prevention

Key Words: RDD, Poststratification, Raking, Variance Inflation.

The 2006 Expert Panel evaluating the BRFSS suggested that the current weighting procedure for BRFSS be improved. Currently, this world's largest random digit dialing telephone survey relies on a simplistic weighting methodology to partially compensate for the bias that results from differential nonresponse and undercoverage. This work examines the current weighting methodology of BRFSS, identifies opportunities for improvements, and introduces an alternative weighting procedure that has greater potential for reducing nonresponse and undercoverage biases without introducing any appreciable inefficiency vis-‡-vis variance inflation. Moreover, a general evaluation methodology is presented for comparing the alternative method against the current weighting procedure, as well as other weighting enhancements designed to compensate for exclusion of nontelephone households.

Diagnostic Process to Assess the Effects of Truncating Extreme BRFSS Sampling Weights

❖Henry Roberts, Centers for Disease Control and Prevention, 4770 Buford Hwy NE MSK66, Atlanta, GA 30341, *hdr9@cdc.gov*; Elizabeth Hughes, Northup Gruman Inc; Robert Woldman, North Carolina Department of Health and Human Services; Ruth Jiles, Centers for Disease Control and Prevention

Key Words: BRFSS, Weighting-Class Variables, Nonresponse Bias, Nonsampling Error, Standard Error

This research focuses on developing and implementing a diagnostic process to study the effects of truncating extreme BRFSS sampling weights; and the extent that extreme sampling weights affects the robustness and precision of BRFSS prevalence estimates. This diagnostic process sequentially assesses: variability, robustness, and precision. The truncation of extreme sampling weights resulted in reducing the variance of the pre-truncation sampling weights for Hispanics and adults ages 18-34 by 298% and 181%, respectively. Estimates that were robust or non-robust prior to the truncation of extreme sampling weights remained robust or non-robust. Overall, the impact of truncating BRFSS sampling weights were: no effect on robustness; large deceases in the variability of the sampling weights; and large decreases in standard errors resulting in enhanced precision.

342 Complex Survey Data Analysis Software GIS ●

Section on Survey Research Methods, Section on Statistical Computing, Section on Statistical Consulting, Section on Statistical Graphics

Tuesday, July 31, 2:00 pm–3:50 pm

Identification of Influential Observations for Regression Using Complex Survey Data

❖Jianzhu Li, University of Maryland, 1218 Lefrak Hall, Joint Program of Survey Methodology, College Park, MD 20742, *jli@survey.umd.edu*; Richard Valliant, University of Maryland

Key Words: influential observations, complex survey, influential observations, cutoff, variance estimations

Even though survey data sets are often large, a few extreme observations can still distort regression estimates. This paper intends to extend the OLS diagnostic statistics such as DFBETAS, DFFITS, and Cook's Distance to the survey setting. Distributional properties and cutoff values will be determined in terms of appropriate variance estimation methods. Extended statistics will be applied to real survey data in a case study.

Goodness-of-Fit Tests for Logistic Regression with Complex Survey Data

❖Amang S. Sukasih, Mathematica Policy Research, Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024-2512, *asukasih@mathematica-mpr.com*; Donsig Jang, Mathematica Policy Research, Inc.; Haixia Xu, Mathematica Policy Research, Inc.

Key Words: unit nonresponse, weighting adjustment, response propensity, logistic regression, simulation

The use of the logistic regression method for unit nonresponse weight adjustments has become common practice in recent years. With this method, users must go through the usual steps in regression modeling, including assessing the goodness-of-fit (GOF) of the model. However, a GOF test that accounts for the complex survey design is not readily available; or if it is, it is not always intuitive. This paper discusses the GOF test for logistic regression with complex survey data. We investigated how much bias the result is when the GOF test for a simple random sample data is applied to data from complex sample design, and whether there is an intuitive pattern in term of bias. We also developed a test that takes into account of the complex survey design. A simulation study was used to compare the simple-random-sample test and the proposed test with readily available GOF tests.

Longitudinal Data Analysis Using SUDAAN

❖Darryl Creel, RTI International, 312 Trotter Farm Drive, Rockville, MD 20850, *dcreel@rti.org*

Key Words: Longitudinal Data Analysis (LDA), SUDAAN

SUDAAN: Software for the Statistical Analysis of Correlated Data (SUDAAN) can be used to analyze data from surveys with complex designs. A possible feature of a complex survey design is clustering. One way in which clustering can occur is to have the same information collected on a sampling unit at different points in time. This type of clustering creates data that may be referred to as longitudinal, panel, or repeated measures data. This paper provides an example of longitudinal data analysis using SUDAAN. The example covers the structure of the data and data set; analytic strategies and interpretation; and the implementation of the analytic strategies using SUDAAN.

Comparison of the Five Test Statistics Currently Available in the SUDAANÆ Modeling Procedures

❖G. Gordon Brown, RTI International, 3040 Cornwallis Rd, PO Box 12194, Research Triangle Park, NC 27709, *ggbrown@rti.org*

Key Words: SUDAAN, Wald F test, adjusted Wald F test, Satterthwaite's adjusted Chi-sq test, Satterthwaite's F test

RTI's software package SUDAANÆ is specifically designed to analyze cluster correlated data or complex data obtained from sample surveys. Currently, SUDAAN offers the choice of five different test statistics for hypothesis testing in the modeling procedures. These test statistics are the Wald Chi, Wald F, adjusted Wald F, Satterthwaite's adjusted Chi, and Satterthwaite's F. It is well known that the Wald Chi test tends to perform poorly when the data pertains to only a few clusters. The other four tests have been developed in an attempt to compensate for this problem and to better handle other intricacies of the data. While the existing research literature addresses merits of these tests, the question of "Which test should I use" remains at large. This work will attempt to answer this question under certain practical scenarios, such as when the design includes 50 or fewer clusters.

Subpopulation Analysis of Survey Data: Statistical Approaches and Software Options

❖ Brady West, University of Michigan, 3554 Rackham Building, Ann Arbor, MI 48109, *bwest@umich.edu*; Steven Heeringa, Institute for Social Research; Pat Berglund, University of Michigan

Key Words: Survey Data Analysis, Statistical Software, Complex Sample Designs, Subpopulation Analysis

In recent years, general-purpose statistical software packages have incorporated new procedures for design-based analysis of complex sample survey data, which feature several useful analysis options. A common and frequently desired technique for analysis of survey data in practice is the restriction of estimation to a particular subpopulation of interest. We discuss two alternative approaches that analysts of complex sample survey data can follow when analyzing subpopulations, and consider the implications of each for estimation and inference. We then present examples of both approaches using data from the National Health and Nutrition Examination Survey (NHANES), and indicate how procedures for survey data analysis in SAS, SPSS, Stata, and SUDAAN respond when following each approach. We conclude with suggestions for practice.

GIS Data Collection for Longitudinal Surveys

❖ Sarah M. Nusser, Iowa State University, Department of Statistics, 222 Snedecor, Ames, IA 50011-1210, *nusser@iastate.edu*

Key Words: CASIC, data collection, agricultural surveys, environmental surveys, geographic data, GIS

Agricultural and environmental surveys typically involve collecting data on a sample of land units. While the foundations of sampling have relied on geographic information systems (GIS) for several years, survey data collection via a GIS interface is relatively rare. A computer-assisted geographic feature data collection instrument was implemented for the USDA's 2005 National Resources Inventory, a longitudinal survey of natural resources that involves repeated observations on area segments. Designing an approach that effectively records, displays and stores polygons that track change over time proved challenging. We will discuss methodological and design considerations in creating a GIS survey instrument based on this experience.

Economic Evaluation of Road Improvements Using GIS and Propensity Score Matching: Case Example from the Republic of Georgia

❖ John Felkner, NORC at the University of Chicago, 1155 East 60th Street, Room 372A, Chicago, IL 60637, *felkner-john@norc.org*; Fritz Scheuren, NORC at the University of Chicago; Celeste Tarricone, Millennium Challenge Corporation; Mamuka Shatirishvili, Millennium Challenge Georgia; Temur Paksashvili, Republic of Georgia

Key Words: GIS, Economic Development, Developing Countries, Access Indices, Propensity Score Matching, Double Differences

The improvement of road and infrastructure networks is crucial for economic development worldwide. The authors present an innovative approach to rigorous statistical evaluation of infrastructure economic impacts by combining indices of city/town travel-times to infrastructure derived from Geographic Information Systems (GIS), with "before" and "after" comparisons using statistical matching techniques and double-differences. The variance between treatment and comparison groups is continuous, while matching and double differences removes selection bias due to the observed differences between treated/comparison, and corrects for possible bias from differences in time-invariant unobserved characteristics. The authors are currently using this approach in the Republic of Georgia, as part of a US $100 million major highway upgrade, sponsored by the US Millennium Challenge Corporation (MCC).

343 Contributed Poster Presentations

Social Statistics Section, Section on Statistical Computing, Section on Statistical Graphics, Section on Statistics in Sports, Section on Survey Research Methods, Section on Quality and Productivity, Section on Risk Analysis, Section on Statistics and the Environment, Section on Statisticians in Defense and National Security, Business and Economics Statistics Section

Tuesday, July 31, 2:00 pm–3:50 pm

Modeling Enrollment Growth at California State University, East Bay

❖ Barbara Warsavage, CSUEB, 516 Soquel Av 4, Santa Cruz, CA 95062, *barbaraw@pacbell.net*

Key Words: Enrollment model, growth model, repeated measures, longitudinal studies

Enrollment growth at a state university is dependent on current enrollment, retention and new enrollment by transfer and by first time freshmen. Prediction of new enrollments can be enhanced using demographic data from the local area. By using these factors in a model known as the partial adjustment model, it is possible to improve projections over standard Markov modeling. A variety of estimation methods will be tried using various error structures. The model will be tested on historical data. Projections will be made with error bounds created by normal theory as well as by bootstrap methods.

Model-Based Clustering Analysis: An Application to Food and Nutrition Data

❖ Jingjing Chen, Student, 531 Williamsburg Wway, King of Prussia, PA 19406, *jjchen_77@yahoo.com*

Key Words: cluster analysis

Cluster analysis is a common technique for the identification of groups of observations that are cohesive and separated from other groups, which is being applied to a variety of areas, such as biology, market research, image analysis, social network analysis, and data mining. Model-based clustering arose from the multivariate normal mixture models with covariances parameterized by eigenvalue decomposition, which is hierarchical agglomeration based on the classification likelihood and the expectation-maximization algorithm for maximum likelihood estimation of multivariate mixture models. The purpose of this study is to review the key concepts on clustering methods and illustrate how to apply various clustering methods in data exploration with an emphasis on model-based clustering. Mike's food and nutrition data was used to demonstrate the application of model-based clustering method.

Geospatial Modeling for Water Chemistry in a Stream Network

❖Stephanie Fitchett, Florida Atlantic University, 5353 Parkside Drive, Jupiter, FL 33458, *sfitchet@fau.edu*

Key Words: spatial models

Findings are presented on the suitability of two distance metrics and several watershed level covariates for geospatial modeling of water chemistry in stream networks. The data are from the Maryland Biological Stream Survey and county level monitoring efforts in 1995–1997.

Predicting Survival from Microarray Data: A Comparative Study

Hege Marie Boevelstad, University of Oslo; ❖Staale Nygaard, University of Oslo, PO Box 1053 Blindern, Oslo, International NO-0316 Norway, *staaln@math.uio.no*

Key Words: Cox regression, Gene expression data, Survival analysis

Survival prediction from high-dimensional genomic data is subject to much research effort in today's medicine. When dealing with these kinds of data, one has to take into account the problem of both censored observations and that the number of genetic variables p usually by far exceeds the number of patients n(p>>n). Most of the proposed methods handle this by combining the Cox proportional hazards model with some estimation regularization technique. We have investigated the performance of seven such methods; two subset selection methods, principal components regression (PCR), supervised PCR, LASSO, ridge regression, and our own modified partial least squares (PLS) Cox regression method. We will describe the estimation techniques and criteria used for evaluating the performance of the methods, and present the results from applications to three published microarray gene expression datasets.

Stalactite Plot for Outlier Detection in the Presence of a Computationally Singular Covariance Matrix

❖Jeremy Nadolski, Benedictine University, 5700 College Rd, Lisle, IL 60532, *jnadolski@ben.edu*; Pablo Marquez, Benedictine University; Lee Ann Smith, Benedictine University

Key Words: outlier detection, stalactite algorithm, singular value decomposition, Mahalanobis distance

Outlier detection is critical in any data analysis project. Outliers affect the ability to predict outcomes successfully and deteriorate accuracy. We used an improved stalactite algorithm proposed by Szychowski et al (2005) to detect outliers in our data. The algorithm is reliant upon Mahalanobis distance which uses the inverse of the covariance matrix. As such, non-invertible matrices in our datasets, which were either noninvertible or computationally non-invertible, could not be supported. To circumvent the non-invertible matrix, we used Singular Value Decomposition (SVD) to create a pseudo-inverse. Using the pseudo-inverse, we were able to replicate the results of Szychowski et al. (2005) and complete our algorithm to find outliers. Our outlier detection algorithm is applied to a dataset on Drosophila melanogaster egg-laying behavior from different wild type strains.

Reckless or Responsible: A Multivariate Statistical Analysis of Consumer Spending

❖Shannon Grant, SUMSRI, 10188 Memory Lane, Rathdrum, ID 83858, *sgrant@uidaho.edu*; Emilola Abayomi, SUMSRI

Key Words: Consumer Spending

As Americans spend more and save less, there is a need to evaluate variables which influence spending habits. First, we reduce the number of variables with principal components analysis and identify underlying factors by grouping correlated variables in factor analysis. Finally, we use discriminant analysis to develop a rule for classifying individual consumers as either reckless or responsible spenders.

SASweave: Literate Programming Using SAS

❖Russell Lenth, The University of Iowa, Department of Statistics, 214 SH, Iowa City, IA 52242, *russell-lenth@uiowa.edu*; Søren Højsgaard, Aarhus University

Key Words: SASweave, Sweave, Literate programming, LaTeX

SASweave comprises software that allows one to embed SAS code in a LaTeX document. After processing the document with SASweave, the SAS output and any graphs produced are added to the document. This guarantees that the output shown is exactly what is produced by the displayed code. SASweave is patterned after Sweave, which does the same thing for R (or S) code. In fact, SASweave incorporates Sweave, allowing both SAS and R statements and output to be embedded in the same document. The presentation will give an overview and several examples of its effective use.

Maximum Likelihood Estimation Using Iterative Importance Sampling

❖Fassil Nebebe, Concordia University, Montreal, QC H3G1M8 Canada, *fnebebe@alcor.concordia.ca*; Tak Mak, Concordia University

Key Words: Monte Carlo methods, Latent variables, importance functions, approximating information matrix

Monte Carlo methods have been a popular alternative to high dimensional numerical integration in maximum likelihood inference involving incomplete data or latent variables. It is well known that the efficiency of Monte Carlo methods can be substantially increased by employing importance sampling. By using a class of importance functions with certain desirable properties, we consider the optimal choice of an importance function in a new setting of iterative importance sampling. We study the statistical properties of the resulting estimator using the present Monte Carlo optimization and examine also the issue of approximating the information matrix.

The Power of Linkage Analysis of a Quantitative Disease Endophenotype

❖Zhuying Huang, Stony Brook University, 69 Eastwood Blvd, Centereach, NY 11720, *zhuying@ams.sunysb.edu*; Nancy Mendell, State University of New York at Stony Brook

Key Words: quantitative trait, power of linkage, pleiotropy, simulation

The power of linkage analysis has been widely studied for simple mendelian diseases. In this report, we consider a complex disease. We assume there is a pleiotropic quantitative trait locus that also affects disease risk. The sample of sib pairs being used consists of a disease-affected proband and plus a random chosen sibling. The linkage analysis is performed on the quantitative trait. We present the results of a power study of this design done through simulation. We consider the cases where theta equals 0.05 and the quantitative trait has high heritability and penetrance values range from 0.01 t0 0.5.

Cluster Visualization in a Two-Dimensional Grid Space by Latent Class Model

❖Chih-Wen OuYoung, National Taiwan University of Science and Technology, 14 Lane 117 ZenAi Rd, Taoyuan 330 Taiwan, *d9401401@mai.ntust.edu.tw*; Yueh Yu[illegible] C[illegible] University of Washington

Key Words: data mining, clustering, latent class model, homogeneity analysis

Clustering analysis has been an useful tool to discover the hidden classes lurking in data with high-dimensionality, especially when they are measured in a discrete format. The challenge part of it not only comes from the decision on the proper clustering algorithm but also on the selection of appropriate data features. Nevertheless, the output classes may sometimes have meaningless structure. Partly to resolve it, researchers apply homogeneity analysis to compress and visualize the high-dimensional data in a two dimensional Euclidean space; however, this approach may be misleading if the true data dimension is more than two. In this paper, we propose a latent class model that superimposes the inter-relationship inherited in a high-dimensional data into a two-dimensional grid space.

Development of Nearest-Neighbor Classifiers Identifying Dermal Sensitizers Based on a Local Lymph Node Assay Database

❖ Shengqiao Li, Centers for Disease Control and Prevention, 1095 Willowdale Road MS L4050, Morgantown, WV 26505, *swl4@cdc.gov*; Adam Fedorowicz, Centers for Disease Control and Prevention

Key Words: Nearest Neighbor Classification, KNN, QSAR, Skin Sensitization, Sensitivity, Mahalanobis Distance

K Nearest Neighbor classifiers were developed to predict skin sensitization of a new chemical based on a murine local lymph node assay database of 178 organic chemicals. Two filters were compared for pre-selection of molecular descriptors. The Fisher's Discriminant Ratio filter picked a subset of descriptors which turn out to be more discriminatory than those picked by the t-test filter. Then, a step forward search method was implemented to screen out extra descriptors and simplify the classifiers based on leave-one-out accuracy. Euclidean and Mahalanobis distance metrics were also examined and the results showed the Mahalanobis distance was appropriate for this study. The 3-nearest neighbor classifier of 13 descriptors singled out by the above methods has an especially balanced performance with sensitivity of 92% and specificity of 81% for this unbalanced dataset.

Methods for K-Models Clustering

❖ Li Li, George Mason University, 4400 University Dr, MSN 6A2, Fairfax, VA 22030, *lli1@gmu.edu*; James E. Gentle, George Mason University

Key Words: data mining, clustering, k-means, simulation

The K-means method is often used in exploratory data analysis to partition data into k groups. The underlying idea is to form groups so as to minimize within-group variation. This is generally effective if the groups are roughly spherical and are separated from each other. If, however, the groups have other structures, such as linear patterns, k-means may not be effective in identifying the groups. Motivated by these considerations, we propose a "K-models" clustering method. We evaluate its performance and study the effect of different starting strategies for K-models clustering.

Netflix Data Mining Competition

❖ Carl Gullans, The College of New Jersey, 3176 Morgan Drive, Wantagh, NY 11793, *gullans2@tcnj.edu*

Key Words: Data Mining, Netflix Competition, Predictive Modeling, Neural Network, Ordinal Response, Decision Tree

Netflix, an online movie rental service, is running a competition to see if their predictive model for customer satisfaction can be improved. Netflix gives $50,000 annually and $1,000,000 in 5 years to the best model. 17,770 files of training data have been provided with each file containing hundreds or thousands of rows, one for each customer rating for that particular movie (scale of 1-5). There are over 100,480,000 total ratings and four columns of information if the files are combined. These four columns are Customer ID, Movie ID, Rating, and Rental Date. Primarily using this data, but with the option to use information such as stars and directors, we will build a model that predicts ratings a set of customers will give to movies that they have not yet seen. We will use various statistical techniques in our model and will explain why certain techniques work better than others.

A New Ecological Model for Insect Population Size with Local Collapse

❖ James Matis, Texas A&M University, Department of Statistics, College Station, TX 77843-3143, *matis@stat.tamu.edu*; Thomas Kiffe, Texas A&M University; Timothy Matis, Texas Tech University; Gerald Michels, Texas A&M University

Key Words: nonlinear regression models

The classic, widely used logistic growth model is based on the mechanistic model N'=(r-sN)N, where N denotes population size, N' its derivative, and r,s>0 birth and death rate coefficients. The positive root, r/s, is called the carrying capacity. Letting F(t) denote the integral of N(t), we develop the new mechanistic model N'=(r-sF)N. The death rate in this model is a function of F, the cumulative past population size, which results in eventual population collapse. The analytical solution of the model may be expressed with parameters having natural physical interpretation. We have fitted the model successfully to numerous data sets on the local abundance of mustard, pecan, cotton, corn leaf and greenbug aphids. We show that a 'model-based' statistical analysis, based on the estimated kinetic rates r and s, has advantages over a standard 'model-free' analysis.

Assessment of Plant Communities Exposed to Fugitive Dust Along a Mine Transportation Corridor in Alaska

❖ Melanie Edwards, Exponent, Inc., 15375 SE 30th Place, Suite 250, Bellevue, WA 98007, *edwardsm@exponent.com*; Elizabeth Maier, Exponent, Inc.; Dawn Reeder, Rare Earth Science; Scott S. Shock, Exponent, Inc.; Nicholas Gard, Exponent, Inc.

Key Words: Vegetation community, Dust impact, Multivariate statistics, Metals

Tundra plant communities in northwestern Alaska were surveyed to characterize the community composition along a haul road from the Red Dog mine to a Chukchi Sea port. Sampling was conducted generally at 10, 100, and 1,000 meters from the road, with 10 microplots per station. The number of species present and percent cover were measured. Collocated soil samples were analyzed to relate to the plant community. Several statistical approaches were used to evaluate the influence of the road on the vegetation community. Results indicated that plant community characteristics vary with distance from the road, with stations furthest from the road similar to their respective reference communities. Metals toxicity as well as physical/chemical stresses typically associated with gravel roads in tundra environments may have contributed to differences observed in plant communities near the road.

Is There a Weekend Effect in Diurnal Temperature Range?

❖ Madhuri S. Mulekar, University of South Alabama, 307 Univ Blvd ILB 325, Mathematics Statistics, Mobile, AL 36688, *mmulekar@jaguar1.usouthal.edu*; Sytske Kimball, University of South Alabama

Key Words: weather, temperature, spectral analysis

Human activities have been linked to the changes in the earth's climate. Studies indicate increase in the global mean surface air temperature during the industrial growth period. Since human activities tend to follow weekly cycles, the existence of weekend effect on the surface air temperatures is suspected. Temperature data from four different stations is used to determine the existence of weekend effect in the diurnal temperature range and to estimate the size of the effect. The sites are located in coastal areas of Mississippi and Alabama in rural locations. Temperature (and other data)

is collected every three seconds but will be averaged to appropriate time intervals in order to detect shorter (i.e., diurnal) and longer (i.e., monthly to annual) term cycles, including those that may be linked to human activity.

Choice of Randomized Design in Large-Scale Education Experiments

❖Brenda Jenney, Arizona State University, 2417 E Whitton Ave, Phoenix, AZ 85016, *brenda.jenney@asu.edu*; Sharon L. Lohr, Arizona State University

Key Words: Experimental Design, Education, Mixed models, Contamination, Randomization

Educational research often studies subjects in naturally clustered groups of classrooms or schools. When designing a randomized experiment to evaluate an intervention directed at teachers, but with effects on teachers and their students, the power for the treatment effect needs to be examined at both levels. If the treatment is applied to clusters, power is usually reduced, however at the same time, this type of design decreases the probability of contamination, which can also reduce the power to detect a treatment effect. Designs that are considered preferable under ideal conditions may not have the highest power when contamination is an issue or when the treatment needs to be evaluated at more than one level. We study the efficiency of several designs for detecting effects at multiple levels under a unified framework, and explore the effects of contamination in the different designs.

Matrix Plots in Support of Visualization of Matrix Characteristics

❖Ferry Butar Butar, Sam Houston State University, Dept of Mathematics and Statistics, Huntsville, TX 77341, *mth_crh@shsu.edu*; Wathurawa A. Bandulasiri, Sam Houston State University; Melinda M. Holt, Sam Houston State University; Cecil Hallum, Sam Houston State University

Key Words: matrices, visualization, plots, pseudoinverse

By its structure, a matrix can easily be visualized graphically. This can be useful (or, at least, interesting) in providing a means to visualize some of its associated characteristics. Numerous matrices result from a convergent sequence of matrices; for example, the pseudoinverse is known to result from a convergent sequence (similar to the sequence used in ridge regression). In such situations, one can "watch" a matrix "go to its pseudoinverse" (i.e., graphically). The singular value decomposition of a symmetric matrix is expressible as a convergent sequence of matrices; again, here one can visualize a matrix converge to its singular value decomposition (or the square root) of a given matrix. This paper investigates a number of such situations whereby visualization are quite insightful. The pros and cons of such an approach will be reported along with accompanying examples.

Statistical Inference in the Presence of Panel Segmentation in Ranking Data

❖Lin Xie, Kansas State University, 101 Dickens Hall, Department of Statistics, Manhattan, KS 66506, *lxie@ksu.edu*; Paul I. Nelson, Kansas State University

Key Words: Panel segmentation, Probability distance model, Ranking data, Parameter estimation, Simulation study

Ranking scales are common in sensory evaluations carried out by a panel of judges. Mallows (1957) developed a two parameter distance probability model for analyzing these data. Segmentation refers to situations where the panel consists of sub-groups of judges whose rankings tend to agree. We use mixtures of Mallows distance distributions to model segmentation. This report develops and explores parameter estimation for these models.

Generalization of the Paired T-Test for the Missing Values Case

❖Sophie Desjardins, University of Québec at Montréal, 7208 Duchesse de Bassano, Montréal, PQ H1N 3X3 Canada, *mcewang@yahoo.com*; Alain Desgagne, University of Québec at Montréal; Pierre Lafaye de Micheaux, Universite Pierre-Mendes France; Benoit Liquet, Universite Pierre-Mendes France

Key Words: T-Test, Missing Values, Paired Data, Comparison of Means, Power

For the comparison of the means of two variables when the observations are paired, we usually use the paired t-test. However, sample data with missing values often occurs and the t-test is not designed to handle them. In this paper, we generalize the t-test to consider the case with missing values and paired observations. A statistic for the test is proposed and it is shown that its distribution is close to a Student-t distribution. The degrees of freedom are also given. Normality of variables is assumed. The power of our test is compared to the one of alternative methods. Applications to real examples are given to illustrate the simplicity of its use.

Use of Beta Regression To Model Adherence to HIV Medication

❖Che Smith, The University of North Carolina at Chapel Hill, 226 Southpoint Crossing Dr, Durham, NC 27713, *ches@email.unc.edu*; Yin Yin, GlaxoSmithKline; Sharon Murray, GlaxoSmithKline; Steven Kathman, GlaxoSmithKline

Key Words: beta, regression, adherence

Knowledge of patient adherence to prescribed medication is of great value in HIV clinical trials to help understand and predict trends in viral load. It is also important to be able to make predictions as early as possible to save time and money. Using data from a phase III trial of 51 patients taking highly active antiretroviral therapy (HAART), we fit and compare several statistical models to represent patient adherence to medication. We fit a beta distribution with parameters estimated using the method of moments. Models were constructed using the CAPABILITY and NLMIXED procedures in SAS, which rely on maximum likelihood estimation of parameters. The method of moments provides the best fit of the adherence data.

A Modified Bonferroni Procedure for Multiple Tests

❖Daniel Mundfrom, University of Northern Colorado, Campus Box 124, 501 20th Street, Greeley, CO 80639, *daniel.mundfrom@unco.edu*; Michelle Roozeboom, University of Northern Colorado; Jamis Perrett, University of Northern Colorado

Key Words: regression, multiple tests, type I error, power

The Bonferroni adjustment ensures that the overall probability of Type I error is below the nominal level when multiple tests are performed. However, the method results in a Type I error rate that is often well below the nominal level. The purpose of this research was to establish an improved single step method for multiple tests based on the Bonferroni method. This method was compared to the Bonferroni and Holm methods to evaluate statistical power while maintaining a Type I error rate that is no greater than the nominal level. Sample data (n = 25, 50, 200) were simulated to have normal distributions for testing of 20 or fewer coefficients. Correlations between x and y and among the xs were varied. Based on the results, a new smaller divisor resulted in increased statistical power while maintaining a level of significance that is closer to, but still less than, the nominal level.

Multivariate Statistical Methods Applied to Digital Photo Processing in Support of Object Detection

❖ Huixing He, Sam Houston State University, 819 15th Street Apt 15, Huntsville, TX 77340, *huixing@shsu.edu*; Cecil Hallum, Sam Houston State University

Key Words: multivariate, detection, image, digitize, object

Retrieving and assessing the information contained in digital image requires more efficient analyses methods. In this study, we develop a capability to permit a "slicing and dicing" of digital images to permit digitizing selected scenes and for conducting specific multivariate analyses in support of the detection of "sought-after" objects. Information enhancements/improvements are investigated in support of detecting "sought-after" objects with emphasis being given to digital images as the primary data source. This is particularly useful when doing automated searching through numerous digital images for objects that, for example, could lead to the discovery of a body, for assessing the impact due to episodic events (e.g., hurricanes, tornadoes, etc.) or highly "sought-after" items by law enforcement or by Homeland Security.

A Method for Estimating Intensity of a Poisson Process

❖ Chih-Hsiang Ho, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, Las Vegas, NV 89154-4020, *chih-hsiang.ho@unlv.edu*; Sandhya Gunti, University of Nevada, Las Vegas; Heng-Wei Cheng, University of Nevada, Las Vegas

Key Words: Intensity function, Poisson process, Time series

Motivated by its vast applications, we investigate ways to estimate the intensity of a Poisson process. In this presentation, we link the traditional homogeneous and nonhomogeneous Poisson processes (NHPP) to the classical time series via a sequence of the empirical recurrence rates (ERR), calculated at equally spaced intervals of time. We consider a computationally simple algorithm to calculate the total area and also the area for the last 10 recurrence rates under the ERR curve. We conclude that the mean function of an NHPP can be estimated from the ERR values. In addition, we argue by simulation, that the algorithm can be implemented to forecast NHPP observations with various forms of intensity function. An intensity correction factor is defined based on the temporal pattern of the targeted point process.

Monitoring Variation in a Multivariate Process with Sample Size Smaller Than the Dimension

❖ John Young, McNeese State University, 1750 Bilbo St, Lake Charles, LA 70601, *johncyoung@bellsouth.net*; Robert L. Mason, Southwest Research Institute; Youn-Min Chou, University of Texas at San Antonio

Key Words: MVSPC, Covariance Matrix

A number of procedures have recently been developed for monitoring changes in the covariance matrix of a multivariate normal process. All of the methodologies that are based on the generalized sample variance require that the sample size, n, be greater than the number of process variables, p, i.e., $n > p$. We introduce a procedure for monitoring changes in the variation of the covariance matrix for a MV normal process based on samples of size $n > 1$ with no restrictions relative to p. We use a control statistic that is a function of the ratio of the determinants of two separate estimates of the covariance matrix, and the distribution of this statistic is based on the distribution of Wilks' likelihood ratio criterion. The proposed statistic is shown to be insensitive to some of the inherent weaknesses associated with using a statistic based on only the generalized variance of a sample.

Development of New Capability Indices for Multivariate Manufacturing Processes

Jen-Nan Pan, National Cheng Kung University; ❖ Chun-Yi Lee, National Cheng Kung University, Department of Statistics, Tainan, Taiwan, *cylee@stat.ncku.edu.tw*

Key Words: multivariate process capability indices, modified engineering tolerance region, multivariate manufacturing process

With the advent of modern technology, manufacturing processes become very sophisticated and merely a single quality characteristic cannot reflect the product quality. Normally, the abnormality of an industrial process is caused by the problems of several interrelated quality characteristics. In order to develop realistic multivariate process capability indices for multivariate manufacturing processes, the modified engineering tolerance region considering the correlation among multiple quality characteristics is proposed. The results of our simulation study show that the new NMCp and NMCpm capability indices perform better than the previous multivariate process capability indices, MCp and MCpm, etc. Finally, two numeric examples further demonstrate that the new NMCp and NMCpm indices can correctly reflect the true performance of a multivariate manufacturing process.

Software Alternatives for Variance Estimation in the Analysis of Complex Sample Surveys: A Comparison of SAS Survey Procedures, SUDAAN, and AM

❖ Bethany Bell-Ellison, University of South Florida, 4202 East Fowler, EDU 162, Tampa, FL 33620, *bbell@coedu.usf.edu*; Jeffrey Kromrey, University of South Florida

Key Words: complex sample surveys, variance estimation, software comparison

The use of standard statistical software for the analysis of data collected from complex sample surveys yields biased point estimates of parameters and incorrect standard errors as a result of unequal probabilities of selection of observations, stratification and clustering, and unit nonresponse. The appropriate analysis of data from complex sample surveys requires the use of software that incorporates sample weights and accurate estimates of variance (e.g., Taylor series approximations or replication estimates). This paper provides a review of the special statistical requirements for the analysis of complex survey data; a comparison of the available statistics and variance estimation strategies provided in SAS, SUDAAN SAS-Callable, and AM; and a contrast of results across these packages. The paper includes samples of code and illustrations of output from each software package.

Performance of First-Order Weighted Random Effects Model Estimators Under Complex Sampling Designs

❖ Yue Jia, Educational Testing Service, Rosedale Road MS 02-T, Princeton, NJ 08541, *yjia@ets.org*; S. Lynne Stokes, Southern Methodist University; Ian Harris, Southern Methodist University

Key Words: weighting, one-way random effects model, estimation bias, sample survey, NAEP

We consider estimation of parameters of random effects models from a sample collected via a complex multi-stage design. We evaluate the first-order weighted method of moments estimators for estimating the grand mean and two variance components. For these estimators, analytic bias expressions are developed and the accuracy of the expressions is evaluated through Monte Carlo simulation. These expressions are then used to examine the impact of sample size, ICC and the sampling design effects. Our study shows that the first-order weighting method effectively reduces the bias for the estimator of mean, and performs well to estimate the variance components for moderate sample size and ICC level. However, for the vari-

ance components estimation, this weighting method should be used with caution for small cluster sizes, particularly with low ICC levels.

Alternative Methods to Compensate for Provider Nonresponse in the National Immunization Survey

❖Phillip Smith, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, MS E62, Atlanta, GA 30333, *psmith3@cdc.gov*; Benjamin Skalland, NORC at the University of Chicago; Robert Montgomery, NORC at the University of Chicago; Sadeq Chowdhury, NORC at the University of Chicago

Key Words: Weighting class adjustment, partial nonresponse

Data is collected in the National Immunization Survey in two phases. In the first phase, households with 19-35 month-old children are identified using a list-assisted RDD survey and data on children's socio-demographic characteristics are collected. At the end of the RDD interview, consent is requested to contact children's vaccination providers to obtain provider-reported vaccination histories. Data from providers may be missing (i) if the household does not give consent to contact vaccination providers, or (ii) if the household gives consent to contact providers but providers do not respond to the mail survey or give inadequate information. Currently, survey weights are adjusted for provider nonresponse using a weighting-class method based on response propensities. This paper explores alternative methods for accounting for provider nonresponse.

Extensions of a Matrix Approach for Comparing Estimate Precision of a Population Total Under a Many-to-Many Frame Structure

Martin Levy, University of Cincinnati; ❖ ZhiYuan Dong, University of Cincinnati, *dongzn@email.uc.edu*

Key Words: sampling, eigensystem, arc-weight estimator, Horvitz-Thomson estimator

We extend our matrix approach for comparing the precision of an estimator of a population total under a many-to-many frame structure to the case of stratified random sampling. We investigate the key issue that arises in this design, which is called crossover, to make it expressible in matrix terms via two methods, one involving decomposition, the other involving a new arc-weight estimator. The stratified case is made tractable by integrating it into our current matrix notation for the SRSWOR case so that we can characterize both the estimates of the population total and their variances under a perfect frame, and using the Horvitz-Thomson and the Arc-Weight estimators under an imperfect frame. We develop code to implement the decomposition of the parent imperfect frame system into closed sub-systems and reproduce the decomposed system visually.

Review of NHIS Public-Design Structures

❖ Van Parsons, National Center for Health Statistics, 3311 Toledo Rd, Rm 3219, Hyattsville, MD 20782, *vparsons@cdc.gov*; Chris Moriarity, National Center for Health Statistics

Key Words: variance estimation, replication, linearization, BRR

The National Health Interview Survey (NHIS) is a multipurpose health survey conducted by the National Center for Health Statistics (NCHS), and is a source of health information for the United States. Traditionally, the complex-design structure of the NHIS is redesigned every decade, and design-based methodologies are developed for the analysis of NHIS data. Historically, public-use microdata files have been released containing simplified survey design structures as compared to the structures available internally to NCHS analysts. These simple structures facilitate design-based analyses using widely available complex-survey software. Furthermore, in recent years design features have been masked to protect confidentiality. In this paper we examine many of the past and present public-use design structures for strengths and weakness and make comparisons with internal design structures.

Estimation of Exposure Point Concentrations for Probabilistic Risk Assessment Using a Three-Dimensional Estimation Technique

❖ James Gleason, Trihydro Corporation, 1252 Commerce Drive, Laramie, WY 82070, *jgleason@trihydro.com*; Heidi Bowen-Morris, Trihydro Corporation

Key Words: Voronoi diagram, Exposure Point Concentration

When estimating human health or ecological risks from exposure to contaminated subsurface soils, a main input is a representative Exposure Point Concentration (EPC). Typically, contaminant concentrations are measured at discrete depths and locations and the EPC is determined statistically. In probabilistic risk assessment, EPC determination is complicated by the observation that environmental data frequently do not follow a definable distribution and the intensity of exposure is a function of depth. Thus there is a need for a method of three-dimensional estimation that is not driven by parametric assumptions. The estimation method proposed builds on the work of Burmaster and Thompson (1996), creating a three-dimensional model using Extruded Voronoi diagrams. The method will be applied to residual lead concentrations at a hypothetical trap and skeet range.

Spatiotemporal Bayesian Maximum Entropy Modeling of Hydrogen Sulfide Concentrations Using Data Collected at Different Observation Time Scales Near Swine Operations

❖ William Allshouse, The University of North Carolina at Chapel Hill, Department of Environmental Sciences and Engineering, *balls@email.unc.edu*; Marc Serre, The University of North Carolina at Chapel Hill; Devon Hall, Rural Empowerment Association for Community Help; Katherine Mills, The University of North Carolina at Chapel Hill; Steve Wing, The University of North Carolina at Chapel Hill

Key Words: hydrogen sulfide, Bayesian Maximum Entropy, CAFO, hog farms, multiple time scales

Hydrogen sulfide (H2S) is an air pollutant produced by industrial hog operations that provides a unique problem for space/time modeling. Spraying hog waste on adjacent fields is believed to be the main source of the compound, creating temporal variation based on the spraying events and spatial variation based on distance to source. Active samplers recorded 15 minute H2S measurements, while passive samplers made a two-week measurement. The difference in measurement duration required that "soft" 15 minute H2S data be created from the passive samplers. The variance of the soft data was derived by accounting for the difference in time scales in the covariance models. The Bayesian Maximum Entropy (BME) method was used to integrate general and site-specific knowledge from the two data types to produce space/time estimates of H2S concentrations.

A Critical Evaluation of 'Moneyball'

❖Mitchell Watnik, California State University, East Bay, 25800 Carlos Bee Blvd, Dept of Statistics, Hayward, CA 94542, *mitchell.watnik@csueastbay.edu*

Key Words: Baseball, Moneyball, Fisher's Exact Test

The popular book "Moneyball" suggested that the Oakland Athletics use statistically-based methods for making draft selections. The book dares readers to compare the Athletics' success under Billy Beane, but offers only anecdotes as evidence. We compare the success of the Athletics' drafts vs. those of other teams to evaluate whether their strategy yields improved results.

Stochastic Modeling of Sports Data

❖Charles N. Lowe, Marshall University, 1 John Marshall Drive, Department of Mathematics, Huntington, WV 25755, *clowe@marshall.edu*; Alfred A. Akinsete, Marshall University

Key Words: Home advantage, Probability, Randomness, Soccer, Steady state, Stochastic matrix

This work considers the statistical analysis of soccer data of series of games played between 1994 and 2005 by teams in the English Premier League. We investigate the home advantage using the number of goals and points scored both at home and away. Extending the results in Akinsete (2004), we obtain the winning, as well as the drawing and losing probabilities of one team against another. We investigate the stochastic behavior and obtain the steady state probabilities of teams as they transit from one state to another either within the state space S = {win, draw, lose}, the weekly (in the case of the last three years), or yearly transitional positions of teams. We adopt the Clarke and Norman's model as a measure of teams' ranking in the league. The randomness in the performance pattern of the teams is studied using the run test.

You (Should Be) Out of Here! Off-the-Field Factors Influencing National Baseball Hall of Fame Selection

❖John Rasp, Stetson University, DIS Dept Unit 8398, 421 N Woodland Blvd, DeLand, FL 32720, *jrasp@stetson.edu*; Nathan Crowder, Stetson University

Key Words: logistic regression, sports statistics

Election to the Hall of Fame is the crowning achievement of any baseball player's career. In theory, selection should be based upon baseball ability and contribution. In practice, other off-the-field factors appear to play a role. We examine the effects of two of these. The first is the extent to which the era in which a player competed (dead ball, post-war, etc.) influences a player's chance of selection. The second is the degree to which market size affects selection probability. (Do players for New York teams profit from the media attention, while otherwise qualified individuals in small markets languish in media obscurity?) We find that both factors significantly influence selection to the Hall of Fame.

Coors Field: Pitchers Graveyard?

❖Jay Schaffer, University of Northern Colorado, Applied Statistics and Research Methods, McKee Hall 530, Greeley, CO 80639, *jay.schaffer@unco.edu*; Raj Chandran, University of Northern Colorado

Key Words: Earned Run Average, ANOVA, Major League Baseball, Elevation

Data were compiled for the 2006 major league baseball season to analyze the effect of elevation on pitching statistics.

An Analysis of Judges' Inconsistency in the Athens 2004 Olympics Artistic Gymnastics Events Using the Multiplicative Interaction Model

❖Wijith Munasinghe, Kansas State University, Department of Statistics, 101 Dickens Hall, Manhattan, KS 66506-0802, *wijith@ksu.edu*; Dallas Johnson, Kansas State University

Key Words: two- way non-replicated experiments, multiplicative interaction model, interactions, AMMI macros

There was a controversial issue in medal standings in the Men's Individual All-Round Artistic Gymnastics event at the Athens 2004 Olympics. The goal of this presentation is to determine whether any of the judges were inconsistent in their scoring when compared to the other judges. The competition data set can be viewed as data from a two-way non-replicated experiments design structure within each individual event. Interaction between judges and gymnasts can be interpreted as inconsistency on the part of the judges. Error to test for interaction cannot be estimated in the usual manner since the absence of replications. One way of handling this problem is to model the data with multiplicative interaction model using AMMI macros recently developed by Lee. These models allow testing for interaction and determining the interaction's structure if such structure exists.

Assessing Ratings Methods for College Hockey Teams

❖Robin Lock, St. Lawrence University, Department of Math CS Stat, Canton, NY 13617, *rlock@stlawu.edu*; Julia Palmateer, St. Lawrence University

Key Words: sports, ratings, Bradley-Terry, hockey

Many collegiate team sports in the United States are organized with multiple leagues where teams play a fairly full schedule of games against league opponents with a sparser schedule of games between teams from different leagues. This often creates difficulties when teams are to be compared across all leagues for national rankings or seeding in post season tournaments. A number of methods have been proposed to calibrate teams while accounting for performance and the quality of opponents. We use Monte Carlo simulations of multiple seasons to compare several existing models, such as raw winning percentage, rating percentage index (RPI), Bradley-Terry (KRACH) and Poisson scoring rates (CHODR), for rating college ice hockey teams.

Measuring Player Contribution in the NBA

❖David Lewin, Macalester College, 1600 Grand Avenue, Saint Paul, MN 55105, *dlewin@macalester.edu*; Vittorio Addona, Macalester College

Key Words: NBA, Sports, Basketball

For an NBA player, a plus-minus rating can be calculated as the difference between the points scored by his team minus the points scored by the opposing team when he is on the court. This is a crude measure of performance since it does not account for the players who share the court with the player in question. We evaluate player performance in the NBA by treating every period between substitutions as a game, and regressing the outcomes of these games on the players who were on the court for that period. Each player is thus a variable in our model, with the intercept representing home court advantage. The data for this work is from the 2005–2006 season, which included roughly 36000 periods. The coefficients obtained from the regression are "adjusted" plus-minus ratings.

Statistical Application in Law: Empirical Investigation on Suspected Drunk Drivers' Data in New Jersey

❖Ayako Kambara, 2640 Route 70 Bldg1A, Manasquan, NJ 08736, *kambara@optonline.net*

Key Words: Statistical Consulting, Law, Drunk Driving, New Jersey

The purpose of this study is to test empirically to what extent the newly introduced alcohol detection instrument is too difficult to blow. Linear probability and binary logit models are estimated using suspected drunk drivers' data in New Jersey. It is known that the Breath Alcohol Concentration (BrAC) is an increasing function of the air volume submitted by the defendant. Without any functions or instructions to alert defendant to stop blowing, the BrAC can keep elevating. Further, the instrument does not set the breath sample criteria based on defendant's biological factors so equality can be an issue. People can falsely be convicted by the instrument's reported result when the instrument is designed too difficult to blow and when the breath sample criteria is set without considering individuals' biological characteristics.

Education by Nation: A Multivariate Statistical Analysis

❖ Anthony Franklin, SUMSRI, 207 Anglewood Dr, Simpsonville, SC 29680, *amfrankl@coastal.edu*; Ashley Brooks, SUMSRI

Key Words: education

We analyze education systems of 64 countries using multivariate statistical techniques such as principal component analysis, factor analysis, and discriminant analysis. Our goal is to classify countries into two populations, one where the educational system of the country is exceptional and the other where the educational system is fair. Reducing the dimensionality of the data set simplifies this process. "Education is our passport to the future, for tomorrow belongs to the people who prepare for it today."– Malcolm X

Multiple Signal Extraction in Seismic Array Data

❖ Joshua Kerr, California State University, East Bay, 25800 Carlos Bee Blvd., Hayward, CA 94542-3004, *joshua.kerr@csueastbay.edu*; Wolfang Polonik, University of California, Davis

Key Words: time series, nonlinear regression, signal

Signal extraction is, and has been, a very important field for quite some time, and for good reason. Upon receiving a seismic reading at a site, the goal is to extricate the signal of interest from the noise-polluted reading attained. This, in and of itself, is a daunting task that has been grappled with. The task mentioned is further compounded when there are multiple signals of interest embedded within the noisy reading. Once separation is established, one tries to garner where the signal came from and at what velocity. An application of this would be to detect if someone detonated a nuclear bomb under ground at the same time an earthquake happened elsewhere. This poster shows some asymptotic results and an example of its use.

Smoothed Moments Variance Estimator for X-11 Seasonal Adjustment

❖ Daniell Toth, Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Avenue NE, Washington, DC 20212-0001, *toth.daniell@bls.gov*; Stuart Scott, Bureau of Labor Statistics

Key Words: seasonal adjustment, X-11, time series, survey error

Pfeffermann's method for estimating variances for X-11 seasonally adjusted series uses the method of moments at a key step. With or without the use of sampling error autocovariances in the method, a decision must be made on the order of an error process. In the former case, negative estimates of error variance can arise in practice. Using X-11 output, this paper models the unknown noise process in order to avoid inappropriate error autocovariances and to try to improve order selection. Comparisons are made to the methods in Pfeffermann (1994) and Pfeffermann and Scott (1997).

A Market Segmentation Analysis of Multivariate Travel Survey Data

Shuang Liu, University of California, Davis; ❖ Peng Wu, University of California, Davis, Institute of Transportation Studies, 404 Russell Park #3, Davis, CA 95616, *penwu@ucdavis.edu*

Key Words: Travel Behavior Survey, Market Segmentation, Telecommuting

This study revisited the San Diego telecommuting survey data using cluster analysis, a commonly used technique for market segmentation studies. This study first divided commuters into several clusters, each of which represented different travel behavior. We then described and interpreted each cluster in terms of its size, mean factor scores, and other pertinent information. Furthermore, we related the cluster membership to various socioeconomic characteristics and other variables of interest, including presence of children, commute distance and time, and preference and choice of telecommuting. This study allowed a better understanding of transportation model choices from a marketing analysis perspective, which would help make more effective policies to promote telecommuting.

344 Wald Lecture I

IMS, General Methodology, Section on Bayesian Statistical Science
Tuesday, July 31, 4:00 pm–5:50 pm

A Review of Surprises Encountered in Bayesian Model Selection

❖ James Berger, Duke University, Inst of Statistics and Decision Science, Durham, NC 27708-0251, *berger@stat.duke.edu*

Key Words: Bayesian, model selection, p-values, frequentist, search strategies, hypothesis testing

When I started looking at hypothesis testing and model selection from a Bayesian perspective many years ago, I thought the following were true: (i) Use of p-values is better than fixed alpha-level testing, since p-values are conditional on the data. (ii) Frequentist and Bayesian testing are incompatible; for instance, Bayes tests do not depend on the stopping rule in sequential settings while frequentist tests do so depend, necessitating 'spending alpha' for looks at the data. (iii) The best single model is the highest posterior probability model. (iv) Finding the best (or a few of the best) models is sufficient. (v) Model selection priors cannot be derived from the data. I no longer think any of these are true and will review why. Included will be a discussion of large model spaces including methods for effective search in, and effective summarization of information from, such spaces.

345 Deming Lecture

Deming Lectureship Committee, The ASA, IMS, ENAR, SSC, WNAR
Tuesday, July 31, 4:00 pm–5:50 pm

A Modern Framework for Enterprise Excellence

❖Douglas C. Montgomery, Arizona State University, Department of Industrial Engineering, Tempe, AZ 85287-5906, *doug.montgomery@asu.edu*

The last three decades have seen significant advances in product and service quality through the use of statistical methods. It is now relatively routine to use statistical techniques in product design and development, optimization and control of manufacturing processes, supply chain management, and nonmanufacturing and service industries. Industrial statistics has emerged as an important branch of statistical science that focuses on this application environment. Yet, as applications of statistics in business and industry have expanded, the need for a broader roadmap to achieve organizational excellence has grown. Implementing such a roadmap involves many challenges, some technical and others managerial and organizational. This presentation focuses on a framework for achieving enterprise excellence and identifies potential approaches to implementation.

346 ASA Presidential Address and Awards

The ASA, ENAR, IMS, SSC, WNAR

Tuesday, July 31, 8:00 pm–9:30 pm

Statistics: Harnessing the Power of Information

❖Mary Ellen Bock, Purdue University, Department of Statistics, West Lafayette, IN 47907-2067, *mbock@stat.purdue.edu*

"Statistics: Harnessing the Power of Information" is the 2007 conference theme which highlights the power of statistics to inform and guide us through our world. Harnessing that power is equally important for our own strategic planning as an organization and a profession. We need data about ourselves when we make decisions about our future directions. We will use this opportunity to examine some evolving facets of our discipline and its practitioners. Topics will include the current demographics and characteristics of statisticians, the changing research frontier for the field, our role in industry, government and education, and our relationships to other disciplines and our communities. Finally, we look to the future and attempt to predict some directions.

347 Section on Risk Analysis Roundtable with Coffee (fee event)

Section on Risk Analysis

Wednesday, August 1, 7:00 am–8:15 am

Assessing Risks in the Environment

❖Edward Boone, Virginia Commonwealth University, Department of SSOR, 1001 W Main St, Richmond, VA 23284, *elboone@vcu.edu*

Key Words: Risk, Environment

When assessing risks in the environmental and ecological settings, there are often many criteria by which we can measure risk. Unlike the business and financial arenas where monetary risk is an easy and obvious quantity to measure, environmental and ecological risks have no obvious quantity. This is an important question when approaching policymakers who will spend money to clean up or protect various ecosystems. Do the risks outweigh the costs to take action? In this roundtable discussion, we will discuss this quandary and brainstorm how to determine a better measure of risk.

348 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education

Wednesday, August 1, 7:00 am–8:15 am

Study Design in Statistics Education

❖Felicity B. Enders, Mayo Clinic, 200 First Street SW, Ha 706, Rochester, MN 55905, *enders.felicity@mayo.edu*

Key Words: Statistics education, Study design, Systematic review

The emerging multidisciplinary field of statistics education has been home to a wide variety of study designs, in part because randomized trials may be difficult, or even unethical, in the classroom. Consequently, the goal of this roundtable is to brainstorm regarding ideal study designs for statistics education. In order to jumpstart the discussion, we will review study designs used in publications from the Journal of Statistics Education, the Statistics Education Research Journal, and selected articles from The American Statistician.

349 Section on Statistical Graphics Roundtable with Coffee (fee event)

Section on Statistical Graphics

Wednesday, August 1, 7:00 am–8:15 am

Introducing Multivariate Statistics Through Graphics and Geometry

❖Daniel Kaplan, Macalester College, Dept of Mathematics, 1600 Grand Avenue, Saint Paul, MN 55116, *kaplan@macalester.edu*

Key Words: graphics, modeling, multivariate, introductory

Real-world problems are almost always multivariate, but most people's mathematical background don't prepare them to work with multivariate statistics. I'll demonstrate a method used at Macalester College to teach introductory-level students the statistical concepts they need to think about multivariate relationships. The method relies on graphics and geometry and makes it much easier to understand in a concrete way seemingly abstract ideas, such as multi-colinearity, and seemingly advanced methods, such as analysis of covariance. The graphics and geometry open up the "black box" of multivariate statistical modeling so nonmathematically oriented people can use the techniques effectively and with understanding.

350 Section on Statistics and the Environment Roundtables with Coffee (fee event)

Section on Statistics and the Environment

Wednesday, August 1, 7:00 am–8:15 am

Multivariate Environmental Health Surveillance

❖Andrew B. Lawson, University of South Carolina, Epidemiology Biostatistics, Columbia, SC 29209, *alawson@gwm.sc.edu*

Key Words: environmental, risk, spatial, health, biomarker, surveillance

In surveillance of environmental risk, there is a need to consider the linkages between various inter-related outcomes. For example, in health risk assessment, we might be interested in monitoring comorbidity or alternative markers of variation in risk. This area is closely related to syndromic surveillance where ancillary variables are used to make early predictions of adverse risk scenarios. In the spatial domain, the task is essentially the monitoring of spatial distributions of risk variables and their surrogates. The similarity to surrogate biomarker endpoints is marked. Hence, multivariate spatial models could be envisioned where prospective modeling will be the main focus. Ancillary time series also would be jointly modeled.

Bayesian Disease-Mapping Methods and Applications: Past, Present, and Future

❖Ying MacNab, University of British Columbia, Room E417 4480 Oak St, Vancouver, BC V6H 3V4, *ymacnab@interchange.ubc.ca*

Key Words: Bayesian disease mapping, Hierarchical models, CAR and MCAR, empirical Bayes and fully Bayesian methods of inferences, Bayesian DALY, Health science and environmental applications

Bayesian disease mapping methodology has seen enormous developments in the past decade. Statistical models and methods of inference have been developed to facilitate univariate, multivariate, and spatiotemporal disease and health outcome analysis and mapping. The intent of this roundtable is to provide a venue for statisticians working on disease mapping to come together and share their reflections and thoughts on this very active area of research, both in terms of the methodological developments and their timely applications in health science and environmental studies, such as cancer and disease epidemiology, health services utilization and outcome evaluation, and population and public health risk assessment, to name a few. It is hoped that statisticians who are experienced or new to disease mapping will find this a worthwhile gathering of both informal and engaged discussions.

351 Section on Statistics in Epidemiology Roundtable with Coffee (fee event)

Section on Statistics in Epidemiology

Wednesday, August 1, 7:00 am–8:15 am

Methods for Missing Data in Survey Response

❖Xiaoming Sheng, University of Utah, 375 Chipeta Way Suite A, Salt Lake City, UT 84108, *xiaoming.sheng@hsc.utah.edu*

Key Words: item response data, survey response, missing data mechanism, bootstrap under imputation, efficiency

Missing data problems persist in many scientific investigations. Various strategies for analyzing missing data have been proposed, but are limited to data on continuous measurements. This roundtable session will focus on some of the available strategies to analyze item response data in survey response. In particular, we will investigate the effects of popular missing data methods on various missing data mechanisms. Large sample behaviors of estimators in a simulation study that evaluates and compares their performance and data from a quality-of-life survey of lung cancer patients will be used to illustrate the utility of these methods. Participants are welcome to talk about their experiences during the discussion.

352 Section on Teaching Statistics in the Health Sciences Roundtables with Coffee (fee event)

Section on Teaching Statistics in the Health Sciences

Wednesday, August 1, 7:00 am–8:15 am

Biostatistics Curriculum for National Institutes of Health (NIH) Training Grants: Training, Expectation, and Outcome

❖Madhu Mazumdar, Cornell University, Weill Medical College, 343 E 74th St Apt 21D, New York, NY 10021, *mam2073@med.cornell.edu*

Key Words: Translations research, Medical device research, Nutritional epidemiology research, % FTE

Many NIH training programs in clinical and translational research emphasize development of courses in biostatistics. (Co-) Mentoring from biostatisticians for the clinical research training project also is encouraged. Biostatistics faculty and staff at academic medical centers often have to develop and/or participate in these programs. The goal of this session is to discuss some of the following issues: 1) How much training time is allocated for biostatistics? 2) What kind of statistical training is needed to be a successful independent researcher? 3) What are the most appropriate text books? 4) What statistical software is most adequate? 5) What % FTE support do biostatisticians derive from training grants? 6) What are the by-products of these mentoring experiences? I will share my opinions on these issues and would like to gain from the experience of other statisticians.

Balancing Teaching and Research: Statistics in the Health Sciences

❖Novie Younger, The University of the West Indies at Mona, Jamaica, P O Box 25580, KIN 3960, Miami, FL 33102-5580, *novie.younger@uwimona.edu.jm*

Key Words: teaching, biostatistics, research, balance, relevant examples

Collaboration with researchers in the health sciences provides for biostatistics teaching sessions that are relevant examples of applied statistics. Such didactic sessions can be highly interactive through discussion of limitations and benefits of various statistical methods. The different audiences—epidemiology, nutrition, and basic medical sciences graduate students and health workers attending short research skills courses or following degree program modules—achieve heightened awareness of the relevance of the various tools to their own career development. Adequate preparation for beneficial teaching may, however, encroach upon time and resources that should be used for research. For this coffee session, we will discuss the possible need to balance the time and resources academic statisticians devote to teaching versus research and ways to achieve this balance.

353 Introductory Overview Lecture 4: How Scientists Can Work Effectively with the Media

The ASA, ENAR, IMS, WNAR, SSC

Wednesday, August 1, 8:30 am–10:20 am

How Scientists Can Work Effectively with the Media

❖Richard Hayes, Union of Concern Scientists, 1707 H Street, NW, Suite 600, Washington, DC 20036, *rhayes@ucsusa.org*

This presentation is based on my book "A Scientists Guide to Talking with the Media". I will describe the challenging intersection of science and media and provide concrete tips and skills to promote accurate and timely coverage of important scientific developments and statistical information. This talk is appropriate for scientists with limited media experience as well as for those wishing to hone their skills.

354 Boosting and Related Methods for Machine Learning ✪

IMS, Section on Physical and Engineering Sciences

Wednesday, August 1, 8:30 am–10:20 am

Fast Boosting Algorithms for Regularized Linear Regression and Classification

❖Jerome H. Friedman, Stanford University, Sequoia 134, Department of Statistics, Stanford, CA 94305-4065, *jhf@stanford.edu*

Regularized regression and classification methods fit a linear model to data, based on some loss criterion, subject to a constraint on the coefficient values. Different forms of the constraint produce different families of solutions whose members are indexed by the constraining value. For example, ridge-regression, subset selection and the lasso use different constraint forms with squared-error loss. For the lasso, a simple boosting strategy has been developed that rapidly computes close approximations to its complete family of solutions. A simple extension to this strategy is presented that can be similarly used with a wide variety of loss criteria and/or constraint forms, including many that are not convex

mboost: A Package for Model-Based Boosting

❖Torsten Hothorn, Institut fuer Medizininformatik, Biometrie und Epidemiologie, Waldstrasse 6, Erlangen, D-91054 Germany, *Torsten.Hothorn@rzmail.uni-erlangen.de*

Key Words: Generalized linear models, Generalized additive models, Gradient boosting , Survival analysis, Variable selection, Software

Classically, boosting or functional gradient descent algorithms for optimizing various empirical risk functions have been implemented using relatively complex base-learners. Recently, boosting algorithms for fitting generalized linear or additive models have been suggested. The key innovation is the application of componentwise linear models or smoothing splines which allows us for a reformulation in terms of classical linear or additive models. In the former case, the regression coefficients can be interpreted in the usual way. Moreover, those boosting algorithms have been demonstrated to be useful for variable selection in high-dimensional situations. We briefly sketch the underlying theory and demonstrate how to actually fit regression models with many covariates in the R system for statistical computing using the `mboost’ add-on package.

Greedy Methods, Weak Learning, and L1 Optimization

❖Tong Zhang, Yahoo! Inc., 135 Oakland Avenue, Tuckahoe, NY 10707, *tzhang@yahoo-inc.com*

Key Words: boosting, greedy algorithm, L1 regularization

I will give a summary of some numerical convergence results for boosting algorithms from greedy optimization perspective and its approximation using weak learning. I will also discuss the relationship to L1 regularization and compare greedy methods to other L1 optimization techniques.

355 Robust Statistical Machine Learning ● ✪

Section on Statisticians in Defense and National Security, Section on Statistical Computing, Section on Physical and Engineering Sciences

Wednesday, August 1, 8:30 am–10:20 am

SigClust: Statistical Significance of Clustering

❖J. Stephen Marron, The University of North Carolina at Chapel Hill, Department of Statisitcs & O.R., Chapel Hill, NC 27519-3260, *marron@email.unc.edu*

Key Words: SigClust, clustering, HDLSS

Clusterings methods provide a powerful tool for the analysis of high-dimension, low sample size datasets (such as microarray data). A fundamental statistical issue is: Which clusters are “really there,” as opposed to being artifacts of the natural sampling variation? SigClust is proposed as a simple and natural approach to this fundamental statistical problem, and its properties are studied.

A Robust Hybrid of Lasso and Ridge Regression

❖Art Owen, Stanford University, Department of Statistics, Sequoia Hall, Stanford, CA 94025, *owen@stat.stanford.edu*

Key Words: Dantzig selector, lars, ridge regression, robustness

A penalty that behaves like lasso for small coefficients and like ridge for large coefficients is developed. This penalty is a reversed Huber function. The penalty is convex. Like the Huber function it requires scaling. The scaling parameter can be incorporated into a criterion that is jointly convex in it and the regression coefficient vector.

Consistent Multicategory Support Vector Machines

Yufeng Liu, The University of North Carolina at Chapel Hill; ❖Ming Yuan, Georgia Institute of Technology, *myuan@isye.gatech.edu*

Key Words: classification, consistency, margin, multicategory

The Support Vector Machine (SVM) has become one of the most popular machine learning techniques in recent years. The success of SVM is mostly due to its elegant margin concept and theory in binary classification. Generalization to the multicategory setting, however, is not trivial. There are a number of different multicategory extensions of SVM in the literature. In this talk, we review several commonly used extensions and explore Fisher consistency of these extensions. A new class of multicategory SVMs are proposed. We develop efficient numerical algorithms to implement the proposed methods. Theoretical properties are explored as well.

Estimating Spatial Covariance Using Penalized Likelihood with Weighted L1 Penalty

❖Zhengyuan Zhu, The University of North Carolina at Chapel Hill, 27599, *zhuz@email.unc.edu*; Yufeng Liu, The University of North Carolina at Chapel Hill

Key Words: LASSO, Choleski Decomposition, Sparsity, Spatial statistics, Spatial autoregression model

In spatial statistics, estimation of large covariance matrix is of great importance because of its role in spatial prediction and design. The classical approach typically assumes that the spatial process is stationary and the covariance function takes some well-known parametric form, and estimates the parameters of the covariance functions using likelihood based methods. When data is available in both space and time, and the spatial stationarity assumption is not reasonable, sample covariance has been used to estimate the spatial covariance matrix. In this paper we study the covariance estimation problem for a class of nonstationary spatial autoregressive model. By exploiting the sparsity structure in the inverse covariance matrix, we show that a LASSO type approach gives improved covariance estimator measured by several criteria.

356 Exploratory Centers for Cheminformatics Research from a Statistician's Point of View ● ✪

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Wednesday, August 1, 8:30 am–10:20 am

Multiple Mechanisms from a High-Throughput Screen

❖ Douglas M. Hawkins, The University of Minnesota, 313 Ford Hall, 224 Church Street S.E., Minneapolis, MN 55455, *dhawkins@umn.edu*; Lexin Li, North Carolina State University

Key Words: Mixtures, Inverse regression, Dimension reduction

Compound libraries in high-throughput screening often comprise groups of compounds operating by different mechanisms, a feature that bedevils methods that fit a single model to the entire collection. Recognizing the different mechanisms holds the potential for substantial improvement in the screening process. Recursive partitioning was shown some 10 years ago to have the potential to recognize this situation, but is conceptually not the best tool. Promising approaches include linear mixture models. Dimension reduction methods based on inverse regression are natural candidates for study as they can identify, not only different mechanisms affecting the mean function, but also those affecting variability. These approaches will be applied to large structure-activity data bases.

Projects in the Michigan Alliance for Cheminformatic Research

❖ Kerby Shedden, University of Michigan, 439 West Hall, 1085 South University, Department of Statistics, Ann Arbor, MI 48109, *kshedden@umich.edu*

Key Words: cheminformatics

I will discuss several research projects in the Michigan Alliance for Chemoinformatic Exploration (MACE). ChemReader is a machine vision approach to extracting chemical structures from diagrams of small molecules appearing in the scientific literature, a key step in annotating chemical databases with literature references. In contrast to other approaches to this problem, ChemReader uses empirically derived chemical intelligence to identify image features that are both accurately resolvable, and highly informative for indexing a specific chemical database (in this case, NCBI's PubChem database). I will also discuss a mathematical transport model that we have developed to characterize large sets of small molecules according to their cell permeability and cellular accumulation properties. Finally, I will discuss ongoing work on multivariate analysis of PubChem biological assay data.

Efficient Cross-Validation for Tuning and Ranking Cheminformatics Models

❖ William J. Welch, University of British Columbia, Department of Statistics, 333-6356 Agricultural Road, Vancouver, BC V6T1Z2 Canada, *will@stat.ubc.ca*; Hui Shen, University of British Columbia

Key Words: Cross validation, Drug discovery, Model selection, Model tuning, Prediction performance

Cross-validation (CV) is widely used for assessing prediction performance of a model, for tuning a model with respect to user-selected parameters, and for selecting a "best" model. Our NIH Exploratory Center for Cheminformatics Research (ECCR) is currently comparing 15 statistical modeling methods and five sets of explanatory variables for modeling a number of biological assays. Among the 15 statistical methods, K-nearest neighbors requires tuning of K, the number of neighbors, for example. By exploiting the pairing that occurs in comparing one method (a given model and explanatory variable set, with specified values for all tuning parameters) with another we are able to achieve high power to discriminate between their two performances. This adaptive and computationally efficient methodology to find the best method will be demonstrated on cheminformatic data sets from PubChem.

357 Application of Hierarchical Models To Study Genetic and Epidemiologic Risk Factors ●

Section on Statistics in Epidemiology, ENAR, Section on Risk Analysis, Biometrics Section, WNAR

Wednesday, August 1, 8:30 am–10:20 am

The Use of Hierarchical Models for Estimating Relative Risks of Individual Genetic Variants

❖ Marinela Capanu, Memorial Sloan-Kettering Cancer Center, E 307 63rd St, 3rd Floor, New York, NY 10021, *capanum@mskcc.org*; Colin B. Begg, Memorial Sloan-Kettering Cancer Center

Key Words: hierarchical models, genetic risk, Gibbs sampling, pseudo-likelihood

Recent technological progress has led to rapid identification and sequencing of large numbers of genetic variants. Studying the associations between these variants and a particular disease is of great importance to epidemiologists in their quest to decipher the disease etiology. Hierarchical modeling is a technique which has been shown to provide more accurate and stable estimates of individual variants, by incorporating exchange of information through the higher levels of the multilevel model. This talk presents recent research on the application and implementation of hierarchical modeling regression to handle these issues using pseudo-likelihood and Gibbs sampling methods. A real data set from a melanoma case-control study is used to illustrate the methods.

Hierarchical Modeling and Stochastic Variable Selection in the Analysis of Multiple SNPs in Genetic Association Studies

❖ David Conti, University of Southern California, 1501 San Pablo Street, ZNI 445, Los Angeles, CA 90089, *dconti@usc.edu*; Juan P. Lewinger, University of Southern California

Key Words: hierarchical modeling, genetic association, haplotype, stochastic variable selection, SNP, Bayesian analysis

With numerous SNPs available within a candidate gene region, one often evaluates both SNP- and haplotype-level associations. In addition to multiple comparison issues, conventional estimation may lead to unstable and biased estimates due to sparse data. Here, we jointly model the SNPs and introduce a modified interaction term to capture the underlying haplotype structure. This analysis estimates both the risk associated with each variant and the importance of phase between pairwise combinations of SNPs. To avoid unstable estimation due to sparse data, we propose a combination hierarchical model and stochastic variable selection procedure to highlight key SNPs and phase terms while incorporating uncertainty in model selection. We demonstrate the performance of this method under various genetic scenarios using simulations and discuss an application to real data.

Multilevel Approaches to Genotype-Phenotype Association: VUS Classification and SNP Blocking

❖Giovanni Parmigiani, Johns Hopkins University, Department of Pathology, Oncology and Biostatistics, 550 North Broadway, Suite 1103, Baltimore, MD 21205, *gp@jhu.edu*; Edwin Iversen, Duke University; Xi Zhou, Cornell University; Daniele Fallin, Johns Hopkins University; Ingo Ruczinski, Johns Hopkins University; Cristian Pattaro, Eurac

Key Words: variants of unknown significance, bayesian classification, hierarchical models, SNPs

In this lecture I will present two applications of multilevel modeling in the area of genetic association studies. The first deals with determining whether the so-called "variants of unknown significance" on known disease genes are themselves disease-related. The second deals with the common case-control design for genome-wide association studies using SNPs. The common theme of the two applications is to use multilevel models to create, at a "lower" level data summaries that capture salient feature of complex data structures, such as families or structured regions in the genome. These parameters are in turn used as discriminating variables in a classification problem for which the lower level groups are the units.

Hierarchical Models for Biases in Observational Studies

❖Sander Greenland, University of California, Los Angeles, Los Angeles, CA 90095-1772, *lesdomes@ucla.edu*

Key Words: Bayesian statistics, hierarchical models, epidemiologic methods, observational data

Ideally, priors for Bayesian analysis should be based on the best available information about model parameters. In observational epidemiology, much prior information is structural, concerning functional dependencies of target parameters on other parameters or variables. This information is captured naturally in a hierarchical (multilevel) model for the observations. In bias modeling, variables determining the parameters include methodological features of the study, and the model is not identified. As a consequence, results are sensitive to the structure as well as to the prior distributions. Thus sensible results demand sensible structural input. An illustration is given in which unknown classification parameters are modeled as a function of the location and spread of a continuous classification criterion. The model is then applied to a study in which the key exposure indicator is latent.

358 Medallion Lecture IV ✪

IMS, General Methodology

Wednesday, August 1, 8:30 am–10:20 am

Quantile Regression Under Censorship

❖Xuming He, University of Illinois, 725 S. Wright Street, 101 Illini Hall, Champaign, IL 61820, *x-he@uiuc.edu*

Key Words: quantile regression, random censoring, convergence

Censoring challenges any regression analysis, including the quantile regression. However, the conditional quantile functions often have better identifibility than the conditional mean function when the response variable is censored. In this talk, we examine the impact of fixed and random censoring on the estimation of quantile regression models. A new approach to doubly censored response will be outlined, and a discussion of existing algorithms on censored quantile regression will be attempted. In particular, we demonstrate that some commonly used censored regression quantile estimators that do not share the spirit of Kaplan-Meier could be poor in efficiency, but a global approach to redistributing the probability mass of censored observations is both challenging and rewarding.

359 Statistical Methodology for the Analysis of Sleep Studies ● ✪

Biometrics Section, WNAR

Wednesday, August 1, 8:30 am–10:20 am

Generalized Linear Mixed Model Analysis of Multistate Sleep Transition Data: The Sleep Heart Health Study

❖Brian S. Caffo, Johns Hopkins Bloomberg School of Public Health, Dept. of Biostatistics, 615 N Wolfe Street, Baltimore, MD 21205, *bcaffo@jhsph.edu*; Bruce Swihart, Johns Hopkins University; Naresh Punjabi, Johns Hopkins Bayview Medical Center

Key Words: GLMM, sleep, transition data, Poisson log-linear models, proportional hazards, observational data

Sleep is a complex behavioral state that has been shown to be essential for maintaining physical and mental function. Recent research has shown that traditional epidemiologic summaries, such as the respiratory disturbance index and summaries of sleep architecture, fail to capture important aspects of sleep. In particular, the rate of transition between sleep states such as rapid eye movement (REM) sleep, non-REM sleep and wake have been shown to have predictive value that these other summaries lack; further refinements consider stages of non-REM sleep. A flexible Poisson generalized linear mixed model is proposed to analyze sleep transition data. This model is shown to synthesize two common complimentary methods for analyzing sleep transition data: multi-state Cox proportional hazards models and log-linear models on marginal sleep transition contingency tables.

Methods for Signal Extraction from EEG Time Series with Application to Large Studies of Sleep

❖Chongzhi Di, Johns Hopkins University, Department of Biostatistics, 615 N. Wolfe St. E3636, Baltimore, MD 21205, *cdi@jhsph.edu*; Ciprian M. Crainiceanu, Johns Hopkins University

Sleep related indices are believed to be causally associated with several diseases, such as: hypertension, sleepiness and cardiovascular disease. A fundamental measure of sleep is obtained by monitoring electric brain signals (EEG) from one or more channels during sleep. EEG time series are then processed in short time windows (5 or 30 seconds) using Fast Fourier Transform (FFT) and filtered using four scientifically relevant bands: alpha, beta, delta, and theta. The fraction of power in each band is estimated in each time window. These estimates are noisy versions of the underlying power spectrum where the noise is a mixture of filtering resistant noise in the original EEG signal measurement. Penalized spline smoothing for each power band of each subject is used to reduce noise. Signal extraction is performed using two complementary strategies.

Semiparametric Event History Models for Analyzing Human Sleep Data

❖Thomas Kneib, Ludwig-Maximilians-Universität München, Ludwigstrasse 33/II, Department of Statistics, Munich, International 80539 Germany, *Thomas.Kneib@stat.uni-muenchen.de*; Ludwig Fahrmeir, University of Munich; Alexander Yassouridis, Max-Planck-Institute for Psychiatry

Key Words: Bayesian smoothing, MCMC, mixed models, multi-state models, penalised splines

Most sleep studies at the Max-Planck-Institute for Psychiatry in Munich focus on sleep structure and its relation to nocturnal hormone secretion or to psychiatric diseases like depression. Raw EEG signals are recorded at several sites during the night together with concentration of certain hormones in the blood measured repeatedly every 10, 20, or 30 minutes. The EEG signals are classified in several stages such as awake, rapid eye movement and states of non-rapid eye movement sleep. We show how flexible semiparametric event history models based on penalized splines can be applied to analyze categorized sleep data. Similar concepts can be used to analyze raw EEG data through penalized spline based signal regression. Both types of models will be formulated in a Bayesian framework with inference based on either penalized likelihood (posterior modes) or MCMC (posterior means).

360 Recent Developments in Macroeconometrics ●

Business and Economics Statistics Section

Wednesday, August 1, 8:30 am–10:20 am

Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market

❖Lutz Kilian, University of Michigan, Department of Economics, 611 Tappan Street, Ann Arbor, MI 48109-1220, *lkilian@umich.edu*

Key Words: Exogeneity, Causality, Identification, Dynamic multipliers, Oil price shocks, Structural VAR model

Given the existence of reverse causality, cause and effect are no longer well defined when relating changes in the real price of oil to macroeconomic outcomes. This creates special challenges for econometric work on oil price shocks. This paper proposes a new method of identifying the exogenous shocks underlying the real price of oil including shifts in the uncertainty about future oil supplies. I quantify the magnitude and timing of these shocks, their dynamic effects on the real price of oil and their relative importance in determining the real price of oil during specific historical episodes. The paper also develops regression methods that allow us to refute the common notion that an increase in the real price of oil has the same effect on the U.S. macro economy regardless of the underlying cause of that increase.

Optimal Impulse Response Function Matching Estimation

❖Barbara Rossi, Duke University, 213 Social Science Building, Durham, NC 27708, *brossi@econ.duke.edu*; Atsushi Inoue, University of British Columbia; Alastair R. Hall, University of Manchester; James Nason, Federal Reserve Bank of Atlanta

Key Words: impulse response, information criterion, DSGE, model estimation, classical minimum distance estimators

We propose a new Impulse Response Function Matching estimator for the parameter of a structural model based on classical Minimum Distance estimation, where the number of impulse responses is selected according to an Information Criterion. The advantages of our procedure are that: (i) it improves the efficiency of the estimates of the model's deep parameters; (ii) it allows the researcher to select the impulse responses that are more informative about the deep parameters. An empirical application to the estimation of representative Dynamic Stochastic General Equilibrium models show that our method can substantially improve inference.

Structural Vector Autoregressions: Theory and Application

❖Dan Waggoner, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309-3904, *dwaggoner@frbatlanta.org*; Tao Zha, Federal Reserve Bank of Atlanta

Key Words: Identification, Efficient Algorithm, Nonlinear Restrictions, Orthonormal Transformation

Identification for simultaneous multiple-equation models is an important issue. Typically, there is no formal way of checking whether a structural VAR model is identified or not. In this paper we develop a new and easily implementable necessary and sufficient condition for the exact identification of a SVAR model. We also develop a sufficient condition for overidentification. Both theorems apply to models with both linear and some nonlinear restrictions on the structural parameters. Moreover, we derive efficient MCMC algorithms to implement sign and long-run restrictions in SVARs. Using our methods, four well-known identification schemes are used to study whether monetary policy has changed in the Euro area since the introduction of the European Monetary Union.

Projection Minimum Distance: An Estimator for Dynamic Models

❖Oscar Jorda, University of California, Davis, Economics Department, One Shields Ave., Davis, CA 95616, *ojorda@ucdavis.edu*

Key Words: impulse response, local projection, minimum distance

This paper introduces an estimator for dynamic models where possibly the dynamics and the variables described therein are incomplete representations of a larger, unknown system. We call this estimator projection minimum distance (PMD) and show that it is consistent and asymptotically normal. Many times, PMD can provide consistent estimates of structural parameters even when the dynamics of the structural model are insufficient to account for the serial correlation of the data or correlation with information omitted from the model. PMD provides an overall specification chi-squared test based on the distance between the impulse responses of the structural model and their semi-parametric estimates from the data. PMD only requires two, simple, least-squares steps and can be generalized to more complex, nonlinear environments.

361 Bayesian Modeling of Incomplete Heterogeneous Data ● ✪

WNAR, ENAR, Section on Bayesian Statistical Science, Biometrics Section

Wednesday, August 1, 8:30 am–10:20 am

Flexible Modeling of the Mean and Covariance Structure for Heterogeneous Incomplete Multiple Longitudinal Markers

❖W. John Boscardin, University of California, Los Angeles, 51-254 Center for Health Sciences, Biostatistics, Mailcode 177220, Los Angeles, CA 90095-1772, *jbosco@ucla.edu*; Hector Lemus, University of California, Los Angeles

Key Words: Kalman filter, Wishart

We model multivariate longitudinal data on multiple subjects using a state space smoothing spline approach. The covariance parameters for the state space model are subject-specific and time-varying so as to allow for arbitrary heterogeneity, but are modeled hierarchically to facilitate borrowing of information across subjects and across time to the extent supported by the data. The performance and applicability of this model is highlighted using intensive care unit data from head trauma patients.

Calculating the Effect of a Randomized Intervention from Panel Data Subject to Noncompliance

❖Siddhartha Chib, Washington University in St. Louis, Olin School of Business, Campus Box 1133, Saint Louis, MO 63130, *chib@wustl.edu*; Liana Jacobi, The University of Melbourne

Key Words: Bayesian inference, Markov chain Monte Carlo, treatment effect, Potential outcomes, predictive distributions

Motivated by the design of an actual experiment, we consider the problem of calculating the effect of a baseline randomized intervention from panel outcomes observed in subsequent periods. The main complication is that subject compliance with the intervention is less than perfect, perhaps due to observed and unobserved confounders. The goal of this study is to explore two alternative ways of addressing this complication in the context of a specific setting, the JOBS II intervention trial which was implemented to test whether an intervention in the form of a job training program can improve mental health and promote re-employment.

Temporal Configuration Analysis: Hidden Markov Models for Heterogeneous Multivariate Longitudinal Data

❖Steven L. Scott, University of Southern California, Bridge Hall 401H, Los Angeles, CA 90089, *sls@usc.edu*

Key Words: hidden Markov model, longitudinal data, model based clustering, panel data

Multivariate longitudinal data can be described using model based clustering. Clusters are defined by a collection of cross sectional models. A subject's path through the data is described by a sequence of cluster membership indicators modeled using a hidden Markov chain. Analysis of such a model proceeds in two steps. First, interpretable labels are mixture components. Second, one uses the posterior distribution of the hidden Markov chain to learn about how individuals under different conditions move through the discrete set of states. This setup allows treatments and interventions to be modeled in terms of the probability of moving subjects to more desirable states. The method can measure temporal inhomogeneity using Bayesian shrinkage methods, and it can describe non-Markov paths by introducing mixtures of latent Markov chains.

Bayesian Modeling of Longitudinal Processes with Dropout and Noncompliance

❖Michael Daniels, University of Florida, Department of Epidemiology and Biostatistics, 207 Griffin Floyd Hall, Gainesville, FL 32611, *mdaniels@stat.ufl.edu*; Xuefeng Liu, Wayne State University

Joint models for the association of a longitudinal binary and continuous process are proposed for situations where their association is of direct interest. The models are parameterized such that the dependence between the two processes is characterized by unconstrained regression coefficients. Bayesian variable selection techniques are used to parsimoniously model these coefficients. An MCMC sampling algorithm is developed for sampling from the posterior distribution. The models are motivated by, and are used for, the analysis of a smoking cessation clinical trial in which an important question of interest was the effect of the (exercise) treatment on the relationship between smoking cessation and weight gain. In this trial, there was both dropout and noncompliance. We propose ways to adjust for both of these factors in our analysis.

362 Variable Selection and Related Statistical Inference ✪

Section on Teaching Statistics in the Health Sciences, Biometrics Section, General Methodology

Wednesday, August 1, 8:30 am–10:20 am

Variable Selection in Semiparametric Regression Modeling

❖Runze Li, The Pennsylvania State University, Department of Statistics, 326 Thomas Building, University Park, PA 16802-2111, *rli@stat.psu.edu*; Hua Liang, University of Rochester

Key Words: Nonconcave penalized likelihood, SCAD, efficient score, local linear regression, partially linear model, varying coefficient models

We propose a class of variable selection procedures for semiparametric regression models using nonconcave penalized likelihood. The proposed procedures are distinguished from the traditional ones in that they delete insignificant variables and estimate the coefficients of significant variables simultaneously. With proper choices of penalty functions and regularization parameters, we establish the asymptotic normality of the resulting estimate, and further demonstrate that the proposed procedures perform as well as an oracle procedure. Semiparametric generalized likelihood ratio test is proposed to select significant variables in the nonparametric component. We investigate the asymptotic behavior of the proposed test and demonstrate its limiting null distribution follows a chi-squared distribution, which is independent of the nuisance parameters.

Selections

❖Chih-Ling Tsai, University of California, Davis, Graduate School of Management, Davis, CA 95616-8609, *cltsai@ucdavis.edu*

Key Words: Dimension reduction, Selection, Shrinkage

In this talk, we will commence by discussing a general paradigm of selections. Subsequently, we will study regression coefficient shrinkage and dimension reduction.

Overall AIC Selection Strategy for Linear Mixed-Effects Models

❖Hua Liang, University of Rochester, Medical Center, Dept. of Biostatistics and Computational Biology, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *hliang@bst.rochester.edu*; Guohua Zou, Chinese Academy of Sciences

Key Words: AIC, Kullback-Leibler information, model selection, profile likelihood, restricted maximum likelihood

The conventional model selection criterion AIC has been parallelly applied to choose candidate models in mixed-effects models by the consideration of marginal likelihood. Its deficiency was recently noticed by Vaida and Blanchard (2005). Correspondingly, the conditional AIC was suggested. We argue that a more overall measure for the difference between the candidate model and true model should be the joint likelihood of data and random effects. In this paper, we develop two joint likelihood-based model selection criteria for linear mixed-effects models. The criteria are the approximately unbiased estimators of the expected Kullback-Leibler information. Simulation studies show that the proposed method outperforms its counterpart:

conditional AIC. Our criteria are also applied to choose the variables in a semi-parametric regression model for a real dataset.

Conditional AIC for Nonlinear Mixed Effects Models

❖Florin Vaida, University of California, San Diego, Division of Biostatistics, 9600 Gilman Drive, MC-0717, La Jolla, CA 92093, *vaida@ucsd.edu*

Key Words: effective degrees of freedom, model selection

In this paper we propose a model selection criterion for nonlinear and generalized linear mixed-effects model (NLME, GLME). The conditional AIC of Vaida and Blanchard (2005) is extended to NLME, using an appropriate definition of the effective degrees of freedom of the model, rho. This rho was proposed for GLME by Lu, Hodges and Carlin (2006). The criterion approximates the conditional Akaike information and the goodness of the approximation depends on the degree of non-linearity of the model. The conditional AIC is useful when the purpose of the model is subject-specific rather than population-level prediction. We use the criterion for model selection in the analysis of data from an ongoing international study of acute HIV infection.

363 New Policies To Facilitate the Tenure Process for Women in Sciences ✪

Committee on Career Development, Caucus for Women in Statistics

Wednesday, August 1, 8:30 am–10:20 am

New Policies to Facilitate the Tenure Process for Women in Sciences

❖Rebecca Doerge, Purdue University, 150 North University Street, Department of Statistics, West Lafayette, IN 47907, *doerge@stat.purdue.edu*; ❖Regina Liu, Rutgers University, Department of Statistics, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, *rliu@stat.rutgers.edu*; ❖Mari Palta, University of Wisconsin-Madison, 610 Walnut Street, Madison, WI 53726, *mpalta@wisc.edu*; ❖Nancy Reid, University of Toronto, Department of Statistics, Toronto, ON M5S3G3 Canada, *reid@utstat.toronto.edu*

Key Words: tenure procedures, women in sciences, academic career in sciences

There have been very important contributions toward the improvement of tenure procedures over the years, for women in sciences. However, reports on how women fare in academia have consistently demonstrated that, while more women graduate with doctoral degrees in the sciences, they lag behind men in achieving tenure and the rank of full professor. For example, the American Association of University Professors reported in 2000-01 that among full-time faculty, women are disproportionately represented at lower ranks and least well represented among full professors. Many reasons have been proposed for the discrepancy, including gender climate and lack of mentoring for women. However, it is increasingly believed that structural issues and the rigidity of the tenure process put women at a disadvantage. Discussants will discuss innovations at their own institutions.

364 Improving the Agriculture Resources Management Survey: Dealing with Complexity in Government Surveys ✪

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 1, 8:30 am–10:20 am

Improving the Agricultural Resources Management Survey: Dealing with Complexity in Government Surveys

❖David Binder, Statistics Canada (Retired), 49 Bertona Street, Napean, ON K2G4G7 Canada, *dbinder49@hotmail.com*; ❖Cynthia Clark, Office for National Statistics, 1 Drummond Gate, London, SW1V 2QQ United Kingdom, *cynthia.clark@ons.gov.uk*; ❖Frederick Conrad, University of Michigan, 426 Thompson Street, Institute for Social Research, Ann Arbor, MI 48106-1248, *fconrad@isr.umich.edu*; ❖Carol C. House, National Agricultural Statistics Service, 1400 Independence Ave SW, Room 5029-A, Mail Stop 2010, Washington, DC 20250-2010, *carol_house@nass.usda.gov*

Key Words: Complex Surveys, Agriculture statistics, Multivariate estimation

Advances in statistical methodology and technology have fostered increased complexity in survey design, collection, and analysis. A panel of the National Academy of Sciences' Committee on National Statistics has recently issued report on the Agricultural Resources Management Survey which discusses aspects of this complex survey, and proposes state-of-the art methods for survey enhancements. Panel members and representatives of the U.S. Department of Agriculture will discuss the report and its implications for complex surveys.

365 Protecting Confidentiality in Public Use Data by Releasing Synthetic Datasets: Experiences with Genuine Applications ●

Committee on Privacy and Confidentiality, Section on Bayesian Statistical Science, Section on Survey Research Methods

Wednesday, August 1, 8:30 am–10:20 am

Protecting Confidentiality in Public Use Data by Releasing Synthetic Datasets: Experiences with Genuine Applications

❖Jerome P. Reiter, Duke University, Box 90251, ISDS, Durham, NC 27708, *jerry@stat.duke.edu*; ❖John Abowd, Cornell University, CISER, Ithaca, NY 14850-2820, *john.abowd@cornell.edu*; ❖Simon Woodcock, Simon Fraser University, *swoodcoc@sfu.ca*; ❖Sam Hawala, U.S. Census Bureau, 4700 Silver Hill Road, Washington, DC 20233, *sam.hawala@census.gov*

Several statisticians in statistical agencies and academia are developing partially synthetic public use data. These comprise the original units in the sample with some values, such as key identifiers or sensitive attributes, replaced with multiple imputations. This panel comprises experts who have led efforts to create partially synthetic datasets for genuine, large surveys.

The panel will discuss their experiences generating synthetic data, focusing on how to implement the approach in practice.

366 Conflict of Interest Concerns in Medical Statistics

Committee on Professional Ethics, ENAR, Biometrics Section, Section on Teaching Statistics in the Health Sciences

Wednesday, August 1, 8:30 am–10:20 am

Conflict of Interest Concerns in Medical Statistics

❖ Peter B. Imrey, The Cleveland Clinic, Department of Quantitative Health Sciences, 9500 Euclid Avenue/Wb4, Cleveland, OH 44195, *imreyp@ccf.org*; ❖ Saskia le Cessie, Leiden University Medical Centre, Einthovenweg, S-5-42, Postbus 9600, 2300 RC Leiden, Netherlands, *cessie@lumc.nl*; ❖ Stephen Ruberg, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *sruberg@lilly.com*; ❖ Lisa LaVange, The University of North Carolina at Chapel Hill, 137 E. Franklin St., Suite 203, Mail Station: CB#8030, Chapel Hill, NC 27514-4145, *lisa_lavange@unc.edu*; ❖ Susan S. Ellenberg, University of Pennsylvania, Center for Clinical Epidemiology and Biostatistics, 423 Guardian Drive, Philadelphia, PA 19104-6021, *sellenbe@cceb.med.upenn.edu*; ❖ Niels Keiding, University of Copenhagen, Department of Biostatistics, Oster Farimagsgade 5, Copenhagen K DK-101, Denmark, *n.keiding@biostat.ku.dk*

Key Words: conflict of interest, clinical trials, data safety monitoring boards, statistical ethics

Medical research is under increasing scrutiny for conflicts-of-interest. With growing dependence of clinical research on private-sector support, suspicions have been heightened by notable instances of highly questionable research practices. Statistical issues are often central to disagreements about medical research design and analysis. Data from clinical studies are assembled and interpreted under statistical oversight, and statistical decisions may have pivotal influence on how studies are perceived. It seems inevitable that conflict-of-interest concerns encompass statistical work. A prominent journal conditions publication of industry studies on external academic replication of primary statistical results. How should the statistical profession view and respond to increasing alarm over potential conflicts of interest? Panelists with varying perspectives will review this problem.

367 Bayesian Applications in Astronomy and Physics ● ✪

Section on Bayesian Statistical Science, ENAR

Wednesday, August 1, 8:30 am–10:20 am

A Dempster-Shafer Bayesian Solution to the Banff A1 Challenge

❖ Paul Edlefsen, Harvard University, One Oxford Street 7th Floor, Statistics Department, Cambridge, MA 02138, *edlefsen@stat.harvard.edu*

Key Words: Dempster-Shafer, Belief Function, High-Energy Physics, Particle Physics, Bayesian Inference, Higgs Particle

We present a Dempster-Shafer Bayesian solution to the particle physics problem of identifying the Higgs particle. The problem, which involves three Poisson distributions, was presented at the July 2006 BIRS workshop Statistical Inference Problems in High Energy Physics and Astronomy as the "Banff A1 Challenge." The D-S solution is an application of the Poisson Dempster-Shafer Model. It provides a Bayesian posterior that does not require priors on the nuisance parameters. In conjunction with the Plausibility Transform, the D-S Bayesian solution is easy to calculate. The results are impressive, demonstrating the power of the technique.

An Exploratory Statistical Study of the Measured Motion of the Guide Star Used in a New Test of Einstein's Universe

❖ Jingchen Liu, Harvard University, 1 Oxford Street, Department of Statistics, Cambridge, MA 02138, *jcliu@stat.harvard.edu*; Xiao-li Meng, Harvard University; Michael Ratner, Harvard-Smithsonian Center for Astrophysics; Irwin Shapiro, Harvard-Smithsonian Center for Astrophysics

General relativity predicts that the phenomenon of "frame-dragging" slowly alters the direction of spin of Earth-orbiting gyroscopes. NASA's Gravity Probe B satellite measured the orientation of such freely falling gyroscopes with respect to a guide star. This star's motion with respect to "fixed points" in the distant universe thus needed to be measured independently. VLBI measurements spread over 14 years yielded 39 positions of this radio-emitting guide star. We use Bayesian methods and a linear regression model of the star's motion to study the effects of various assumptions involving the parameters of both the physical model and the measurement noise model. We find that a wide range of different assumptions lead to a narrow range of point estimates and error distributions.

Upper Limits for Source Detection in the Three-Poisson Model

❖ Paul Baines, Harvard University, 17 Dimick Street, Somerville, MA 02143, *pdbaines@fas.harvard.edu*

Key Words: Astrostatistics, Bayesian modelling, Coverage properties, Empirical Bayes, Simulation

The Large Hadron Collider (LHC) at CERN is set to produce data that may shed light on the existence, or otherwise, of the Higgs-Boson particles. One type of data to be collected may be modeled using a system of three Poisson models corresponding to experimental observation, in addition to instrumental and background calibrations. Several Bayesian approaches are explored to provide marginal posterior percentiles for the parameter of interest. With multiple decay channels, dimensionality of the nuisance parameter grows and specification of the prior structure becomes increasingly challenging. Problems with vague priors in high-dimensions are illustrated in the multichannel case. A systematic approach is taken using single-level and hierarchical models, with full Bayes and Empirical Bayes methods compared. The goal is to provide upper limits with excellent Frequentist coverage properties.

A Statistical Approach to Stellar Archaeology

❖ Hyunsook Lee, Harvard-Smithsonian Center for Astrophysics, 60 Garden St MS 70, Cambridge, MA 02138, *hlee@cfa.harvard.edu*; Vinay L. Kashyap, Harvard-Smithsonian Center for Astrophysics; Andreas Zezas, Harvard-Smithsonian Center for Astrophysics

Key Words: Isochrones, Color-Maginitude Diagram (CMD), Error Magnitude, Likelihood, Information, Posterior Distribution

In this presentation, we propose a statistical color magnitude diagram (CMD) fitting procedure, which for the first time provides stellar age uncertainty, based on photometric and theoretical model uncertainties. Fitting the theoretical isochrones to color-magnitude data has been recognized as a tool for identifying the age of stars and stellar associations, where the age information is imprinted on the {\it p}-dimensional distribution of their {\it p}-band photometric data. We begin with constructing the likelihood of a star cluster which provides the most probable age of the cluster. The density form of this likelihood, then, proffers the posterior distribution of each individual star that accounts for the age uncertainty due to the photometric measurement uncertainties and astrophysical model uncertainties. Our method is illustrated with stars of the Small Magellanic Cloud.

Coherent Bayesian Inference on Compact Binary Inspirals Using a Network of Interferometric Gravitational Wave Detectors

❖Christian Röver, The University of Auckland, Department of Statistics, Private Bag 92019, Auckland, New Zealand, *christian@stat.auckland.ac.nz*; Renate Meyer, The University of Auckland; Nelson Christensen, Carleton College

Key Words: gravitational waves, coherent parameter estimation

Presented in this paper is the description of a Markov chain Monte Carlo (MCMC) routine for conducting coherent parameter estimation for interferometric gravitational wave observations of an inspiral of binary compact objects using multiple detectors. Data from several interferometers are processed, and all nine parameters (ignoring spin) associated with the binary system are inferred, including the distance to the source, the masses, and the location on the sky. The data is matched with time-domain inspiral templates that are 2.5 post-Newtonian (PN) in phase and 2.0 PN in amplitude. We designed and tuned an MCMC sampler so that it is able to efficiently find the posterior mode(s) in the parameter space and perform the stochastic integration necessary for inference within a Bayesian framework. Examples are given for simulated signals and data as seen by the LIGO and Virgo detectors.

368 Keepers of the Fire ● ✪

Social Statistics Section

Wednesday, August 1, 8:30 am–10:20 am

Institutional Memory and Information Management

❖Paul Johanis, Statistics Canada, *paul.johanis@statcan.ca*

Key Words: information management, institutional memory

Statistical organizations are at risk of losing corporate memory as a result of demographics, changes in organizational behaviors and in particular the advent of electronic information. In this context, the effective management of information holdings becomes a strategic element in capturing and sharing institutional memory. This paper proposes a classification of broad categories of information that are relevant for information management in a national statistical institute and provides examples of the application of this framework in Statistics Canada.

Transforming Statistical Knowledge into Institutional Memory

❖Olivia Blum, Israeli Central Bureau of Statistics, P.O. Box 34525, Givat Shaul, Jerusalem, International 91342 Israel, *blum@cbs.gov.il*

Key Words: Institutional Memory, Human Capital, Social Capital

The search for a mechanism to create and preserve institutional memory emerges when accumulated skills, knowledge, expertise, and professional insight are about to leave an active organization. The use of institutional memory is functional to the mere existence of the organization in a competitive open market, when striving for efficiency under limited resources. Institutional memory is also instrumental to the individuals working in the organization. It generates minimal routine needed to avoid constant decision-making processes. The synergy between the people and the organization brings about traditions in the activation of past processes in knowledge-based rituals. In national statistics institutions the institutional memory is contained as an asset in the close environment of the individual, the organization within the government, and the organization in the international arena.

The Planning Database: Decennial Census Data for Historical, Real-Time, and Prospective Analysis

❖J. Gregory Robinson, U.S. Census Bureau, Population Division Room 5H491, Washington, DC 20233-8800, *j.gregory.robinson@census.gov*; Carrie L. Johanson, U.S. Census Bureau; Antonio Bruce, U.S. Census Bureau

Key Words: Database, Census 2000, demographic variables, targeting, non-response

The Planning Database (PDB) developed for Census 2000 assembled a range of housing, demographic, and socioeconomic variables at the tract level that are correlated with nonresponse and undercounting. The database provided a systematic way to identify potentially difficult-to-enumerate areas to flag for special attention in Census 2000. It was also successfully used for other Census 2000 planning, operational and evaluation activities. This paper reviews the development of the PDB, describes key uses in Census 2000, and outlines more integrated uses of the PDB for 2010 census activities and for ongoing surveys. Some enhancements include a mapping interface, additional data sources, levels of Census geography in addition to the tract level and interactive reporting. We give a case study example of the use of the PDB to target areas with potential low mail response rates in 2010.

369 Application and Issues of Crossover and Matched Design in Clinical Trials ● ✪

Biopharmaceutical Section, ENAR, WNAR

Wednesday, August 1, 8:30 am–10:20 am

A Study of Bioequivalence Trials with Missing Data

❖Yiyun Zhang, The Pennsylvania State University, 325 Thomas Building, Department of Statistics, University Park, PA 16802, *yuz115@psu.edu*

Key Words: Bioequivalence, Crossover design, Missing data, Clinical trials

In a bioequivalence trial, missing data are very likely to occur for various reasons. Traditional treatment for missing data is simply dropping those data with missing and performs tests on the rest of the data. Obviously this throws away valuable information and is not efficient. It is of interest how the missing data problem affects the results, e.g., the power of the tests, etc. Many methods under different settings have been proposed in the literature, for example in Lee et al. (2005) and Lim et al. (2005), etc. Simulation studies are performed using SAS PROC IML for both 2 by 2 and 3 by 3 designs.

Crossover Designs Based on Type I Orthogonal Arrays for a Self and Simple Mixed Carryover Effects Model with Correlated Errors

❖Samad Hedayat, University of Illinois at Chicago, Dept of Math Stat Compt Sci, 851 South Morgan Street, Chicago, IL 60607-7045, *hedayat@uic.edu*; Zhiwu Yan, University of Illinois at Chicago

Key Words: crossover designs, self carryover effects, simple mixed carryover effects, universal optimality, type I orthogonal arrays

We investigate the performance of crossover designs based on type I orthogonal arrays for a self and simple mixed carryover effects model in the presence of correlated errors. Assuming that the between-subject errors are independent while the within-subject errors behave according to the stationary first-order autoregressive and moving average processes, analytical optimality results for 3-period designs are establish and, as an illustration, numerical details for a number of 4-period cases are tabulated.

Uniformly Balanced Repeated Measurements Designs in the Presence of Subject Dropout

❖ Dibyen Majumdar, University of Illinois at Chicago, MSCS mc249, 851 South Morgan, Chicago, IL 60607, *dibyen@uic.edu*

Key Words: Crossover studies, Optimal Designs, Design Efficiency, Connectedness

Uniformly balanced repeated measurements designs have been extensively studied in the literature. It has been shown that they possess elegant combinatorial properties. They are highly efficient crossover designs. In 1999, Low, Lewis and Prescott showed that a crossover design based on a Williams Latin square (which is a uniformly balanced repeated measurements design) of order 4 can suffer substantial loss of efficiency if some observations in the final period are unavailable. Indeed, if all observations are missing, the design becomes disconnected. In this talk we will examine some properties of these designs when there is a possibility that observations may be unavailable due to subject dropout. We also investigate designs with minimal loss of efficiency under subject dropout.

A Monte Carlo Study for the Methods Used for Assessing Treatment Differences with Possible Carryover Effect in Crossover Designed Studies

❖ Ling Chen, Food and Drug Administration, 10903 New Hampshire Ave., WO22, Room 5241, Silver Spring, MD 20993, *Ling.chen@fda.hhs.gov*; Ferrin D. Harrison, Food and Drug Administration

Key Words: carryover effect, crossover design, Monte Carlo study, treatment difference

Chen & Ferrin (2006) proposed a test for assessing treatment differences adjusted by possible carryover effects in crossover designed studies. A Monte Carlo study was conducted to compare the new methodology to the generalized least square approach with respect to the coverage and width of confidence intervals for the difference in treatment means with constant or estimated ratio of between over within subject variance in a crossover study. Results and comments from the study will be presented.

Sample Size Re-estimation for Noninferior Study of Matched Data

❖ Yi Tsong, Food and Drug Administration, Room 5244 Bldg 22 10903 New Hampshire Av, Silver Spring, MD 20993, *tsong@cder.fda.gov*; Wu Yute, Food and Drug Administration

Key Words: non-inferiority, crossover trial, correlation

Paired and matched data have often been used in clinical trials and epidemiological safety studies. For example, crossover and William's square designs are often used in generic drug evaluation, QTc trial, pain relieve drug evaluation and in drug abuse potential trials. Two treatments may be paired on matched to study the treatment efficacy or to perform a case-control safety study. Either the response is continuous or binary variable, the correlation existed between the matched responses is not negligible. Conventionally, the sample size is determined inefficiently using the independent information of the two groups. On the other hand, when the trial is of non-inferiority or equivalence in nature, sample size is determined by a alternative hypothesis of a fixed treatment difference between test and the active control, without prior information of the true difference.

370 Experimental Design and Analysis of Drug Combinations ● ✪

Biopharmaceutical Section, ENAR, WNAR

Wednesday, August 1, 8:30 am–10:20 am

A Hierarchical Paradigm for Synergy Assessment Modeling Approaches

❖ William Greco, Roswell Park Cancer Institute, Cancer Prevention, Elm and Carlton Streets, Buffalo, NY 14263, *william.greco@roswellpark.org*

Key Words: synergy, response surface, interaction, combination, isobol, Hill model

Our group has collected a set of approaches for assessing the nature and intensity of agent interaction in binary and higher order chemical mixtures, and has created a paradigm for applying these approaches in a hierarchical manner from least complex (most model-independent) to most complex (most model-dependent). The approaches, in order, include: (a) fitting the 4-parameter Hill model (or other appropriate model) to concentration-effect data from each individual agent and from each fixed-ratio binary or higher order mixture; (b) constructing isobols at several effect levels; (c) repeating approach (a), but after transforming the concentration variable into the combination index at the 50% effect level; (d) fitting a simple response surface to all of the data at once; (e) fitting a relatively complex comprehensive response surface to all of the data at once. Supported by NIH RR10742.

Nonlinear Blending: An Alternative to Dosewise Additivity, the Interaction Index, and Bliss Independence

John Peterson, GlaxoSmithKline; ❖ Steven J. Novick, GlaxoSmithKline, Mail code: MAI.T0201, Research Triangle Park, NC , *steven.j.novick@gsk.com*

Key Words: combination drug study, mixture experiments, drug synergy

Many classical synergy measures were derived under somewhat idealized pharmacological situations. These measures are rather limited relative to the wide variety of response surfaces that occur in practice. The statistical area of "mixture experiments", however, makes use of a concept called nonlinear blending to quantify synergy. Nonlinear blending by its simple nature is well defined for any shaped dose response surface. Drugs with different relative potencies, different effect maxima, or situations of potentiation or coalism pose no problem for nonlinear blending as a way to assess the increased response benefit to be gained by combining two drugs. This paper introduces for the first time the concept dichotomy of weak and strong nonlinear blending, and shows how strong nonlinear blending can be used for determining whether or not to blend compounds for enhanced efficacy.

Experimental Design and Synergy Analysis for Drug Combination Studies Based on Log-Linear Dose Responses

❖ Hong-Bin Fang, University of Maryland Greenebaum Cancer Center, HSF II S203B, 20 Penn Street, Baltimore, MD 21201, *hfang@umm.edu*; Ming T. Tan, University of Maryland Greenebaum Cancer Center

Key Words: Additive action, Combination Index, Dose-effect, Experimental design, Synergy, Nonparametric model.

We propose experimental designs (dose-finding and sample size determination) for detecting departures from additivity of two drugs with log-linear dose responses in a general statistical model that is applicable to both in vivo and in vitro experiments. The design is derived by means of uniform measures that maximize the minimum power of the F-test to detect any departures from additive action, and at the same time minimizes the maximum bias due to lack of fit among all potential departures of a given meaningful magnitude. Furthermore, we propose a nonparametric combination index model to capture the interaction of two drugs. The nonparametric function is estimated using the technique developed in thin plate splines. A study on two anticancer drugs is used to illustrate the method.

A Semiparametric Response Surface Model for Assessing Drug Interaction

❖ Maiying Kong, University of Louisville, Bioinformatics and Biostatistics, 555 S Floyd Street Ste 4026, Louisville, KY 40202, *maiying.kong@louisville.edu*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: drug interaction, synergy, semiparametric model, bootstrap method

Existing response surface models for assessing drug interactions are not adequate to capture the complex patterns when synergy, additivity, or antagonism exist for two drug combinations. We propose a two-component semiparametric response surface model with a parametric function to describe the additive effect of a combination dose and a nonparametric function to capture the departure from the additive effect. The nonparametric function is estimated using the technique developed in thin plate splines, and the pointwise confidence interval for this function is constructed based on bootstrap method. The proposed semiparametric model offers an effective way of formulating the additive effect while allowing the flexibility of modeling a departure from additivity. Example and simulations are given to illustrate that the proposed model performs very well.

371 Current Issues in Risk Prediction Modeling ●

ENAR, Biometrics Section

Wednesday, August 1, 8:30 am–10:20 am

Evaluating the Incremental Impact of a Risk Prediction Marker Using Predictiveness Curves

❖ Margaret Pepe, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N M2B500, Seattle, WA 98109, *mspepe@u.washington.edu*; Wen Gu, University of Washington; Ying Huang, University of Washington; Ziding Feng, Fred Hutchinson Cancer Research Center

Key Words: predictiveness curve, risk prediction, baseline risk model, perceived risk, marker, cancer

The predictiveness curve displays distribution of risk when a risk model is applied to a population of individuals. Individuals can more easily make decisions when perceived risk is high or low, so models in which most individuals are classified as high or low risk are preferable to models classifying subjects in an intermediate risk range. We describe several methods for comparing these curves. Standard summary measures of predictability (R-squared) are shown to be summary indices for the predictiveness curve. We provide a simple interpretation for R-squared and show how formal comparisons can be based on it. Alternatively, if threshold values are specified for defining high or low risk, comparisons can be based on fractions of the population in these risk categories. We address the problem of comparing a baseline risk model to a one augmented by addition of a risk prediction marker.

C-Statistics and Reclassification in Risk Prediction Models

❖ Nancy Cook, Brigham and Women's Hospital, Division of Preventive Medicine, 900 Commonwealth Ave East, Boston, MA 02215, *ncook@rics.bwh.harvard.edu*

Key Words: prediction, risk factors, ROC curves, discrimination, calibration, cardiovascular disease

The c-statistic, or area under the ROC curve, is popular in diagnostic testing, where the goal is to discriminate diseased from non-diseased patients. The c-statistic may not be optimal for assessing models predicting future risk or stratifying individuals into risk categories, where calibration is as important in assessing risk accurately. Perfectly calibrated models can, in fact, usually only achieve values for the c-statistic well below 1. Accepted cardiovascular risk factors, such as lipids, hypertension and smoking, individually have minimal impact on the c-statistic, yet can more accurately reclassify large proportions of patients into higher or lower clinical risk categories. The joint distributions of risk predictors, along with the calibration of risk within cross-classified categories, can help determine which models more accurately predict risk for individual subjects.

Updating Clinical Prediction Models: Simple Methods Give Promising Results

❖ Kristel J.M. Janssen, Julius Center UMC Utrecht, PO BOX 85500, Utrecht, 3508, Netherlands, *k.j.m.janssen@umcutrecht.nl*; Karel G. Moons, Julius Center UMC Utrecht; Yvonne Vergouwe, Julius Center UMC Utrecht; Diederick E. Grobbee, Julius Center UMC Utrecht

Key Words: Prediction research, Updating, Calibration, Discrimination, Prediction model

When prediction models show poor performance in new patients, one often simply redevelop new rule ignoring prior data. We compared five methods to update a model with new data. The model was developed in 1944 pts, validated+updated in 752 pts, tested in 283 pts. In updating method 1: only intercept adjusted; 2: intercept adjusted+regression coefficients multiplied with calibration slope; 3: tested if predictors were different; 4: re-estimated intercept+regression coefficients in validation set; 5: same as 4 in combined derivation+validation set. Calibration original model in validation set substantially improved by all methods. Discrimination was not improved with method 1+2. In test set all updated rules had similar discrimination. Simple recalibration methods improved model calibration similar as extensive revision methods. Discrimination was similar across all updating methods.

Rehabilitation of the Case-Control Design in Prediction Research

❖ Karel G. Moons, Julius Center UMC Utrecht, PO Box 85500, Utrecht, 3508 GA Netherlands, *k.g.m.moons@umcutrecht.nl*; Cornelis Biesheuvel, Julius Center UMC Utrecht; Yvonne Vergouwe, Julius Center UMC Utrecht; Ruud Oudega, Julius Center UMC Utrecht; Arno W. Hoes, Julius Center UMC Utrecht; Diederick E. Grobbee, Julius Center UMC Utrecht

Key Words: Prediction research, Absolute disease probabilties, Nested case control, Diagnostic research, Prognostic research, Valid and efficient

Case-control studies are widely disapproved for (diagnostic and prognostic) prediction research; potential for selection bias and cannot obtain absolute disease probabilities. Both can be solved with nested case-control design. Using data from diagnostic cohort study, we drew nested case-control samples including all cases. Controls were randomly sampled with ratios 1:1, 1:2, 1:3 and 1:4. Per ratio, 100 samples. We weighted controls by corresponding 1/sample fraction. Predictive accuracy of index tests, including post test probabilities, in nested case-control samples were same as in full cohort. Width of 95% CIs decreased with increasing control number.

Nested case-control is a valid and efficient design for prediction studies. It should be (re)appraised and is notably advantageous if disease is rare, index tests are costly/invasive, or when stored data or human material is reanalyzed.

372 Statistics for Nonlinear Dynamics ✪

IMS, Business and Economics Statistics Section
Wednesday, August 1, 8:30 am–10:20 am

Collocation Tempering for Differential Equation Models

❖David A. Campbell, Simon Fraser University Surrey, 250-13450 102nd Ave, Surrey, BC V3T 0A3 Canada, *campbell@math.mcgill.ca*

Key Words: MCMC, Differential Equation Models, Smoothing, tempering

Differential equation models often have no analytic solution so statistical models centered on the solution to a differential equation require a numerical solution 'S.' Computational Bayesian methods for differential equation models based on MCMC require S to obtain the likelihood. However, a single differential equation can be used to model unbounded growth, limit cycles or a tendency towards an asymptote simply by changing the parameters. Consequently results of MCMC based on the numeric solution are highly dependent on parameter values used to initialize the algorithm because in crossing the parameter space a mediocre fit to the data must deteriorate substantially before it can improve. I present a collocation based tempering method which overcomes these problems while converging quickly to a stable distribution.

Meta-Analysis Based on Summarized Data with Application to Drug-Drug Interaction

❖Lang Li, Indiana University, 410 West 10th street, Suite 3000, Indianapolis, IN 46202, *lali@iupui.edu*

Key Words: Bayesian Model, pharmacokinetics, drug-drug interaction, meta-analysis, Monte-Carlo Markov Chain

An innovative DDI prediction method based on a three-level hierarchical Bayesian meta-analysis model is developed. The first level model (study-specific) recovers study specific PK parameters and between subject variations. The second level model (between-study specific) describes the between study variations, and the third level model (prior distributions) summarizes the prior knowledge of PK parameters and their variances. Seven ketoconazole and five midazolam studies are analyzed through this model, and their in-vivo DDI (measure with area under the concentration curve ratio) is predicted. Monte Carlo Markov chain (MCMC) is used for generating posterior samples of all parameters and DDI predictions. The performance of this model is validated through the statistical simulations.

Detecting Nonlinear Dynamics in a Psychoacoustics Experiment

❖Theodoro Koulis, JSS Medical Research, 4492 Rue Ste-Catherine Ouest, Westmount, QC H3Z 1R7 Canada, *tkoulis@jssresearch.com*

Key Words: functional data, differential equation, mixed-effects, perception-action, psychophysics

Experiments in cognitive science are evolving in complexity in order to answer more difficult and fundamental questions about the mind. As a result, rich data is being collected in real-time for the purpose of exploring the dynamic nature of the brain. For example, continuous response digital interfaces are increasingly used to measure perception-action systems in psychophysics. The statistical problem related to these systems reduces to linking time-varying covariates to a continuous response variate. Often, the data from experiments suggest that an underlying system is linear. We will introduce ideas from functional data analysis, differential equations and mixed-effects modeling, which as a collection of tools, are useful for detecting nonlinear dynamics in perception-action systems.

Testing for Missing Dynamical Variables

❖Giles Hooker, Cornell University, Dept Bio Stat and Comp Bio, Ithaca, NY 14850, *giles.hooker@cornell.edu*

Key Words: dynamics, tests, goodness of fit, chaotic data analysis, differential equations, latent variables

This talk investigates model building in dynamical systems. In many contexts, systems of nonlinear differential equations have been proposed to model complex behavior. These equations have had remarkable success at mimicking the qualitative features of observed systems. However, they frequently exhibit poor quantitative performance. I investigate diagnostics for nonlinear differential equations. In particular, I show that the estimation of lack of fit can be appropriately measured in terms of latent inputs into the dynamical system. In this talk, I discuss the analysis of such inputs for the discovery of missing dynamical variables. The machinery of Chaotic Data Analysis may be employed here to derive a diagnostic for missing components. However, results from dynamical systems theory also suggest that care should be taken when modifying systems to account for such discoveries.

Statistical Data Assimilation for Marine Ecological Prediction

❖Michael Dowd, Dalhousie University, Halifax, NS B3H 4J1 Canada, *mdowd@mathstat.dal.ca*

Key Words: nonlinear dynamics, Monte Carlo methods, Bayesian statistics, state space models, oceanography, ecosystem models

Ocean prediction relies on dynamical models for biological and physical processes, and available time series oceanographic measurements. In this presentation, I overview statistical approaches for combining observations with nonlinear differential equation based mathematical models in the marine environmental sciences. This problem known as data assimilation and has the goal of joint estimation of the time evolving system state, as well as static parameters. The approach taken relies on the nonlinear and non-Gaussian state space modeling framework. Solution techniques involve Monte Carlo methods to approximate the target densities and to treat filtering, smoothing and likelihood based parameter estimation. These ideas are considered for operational ocean prediction using a dynamical marine ecosystem model and near real-time observations from a coastal ocean observing system.

373 Bayesian Statistics in Medical Device Clinical Trials ● ✪

Section on Bayesian Statistical Science, Biopharmaceutical Section, ENAR
Wednesday, August 1, 8:30 am–10:20 am

Weighing Evidence and Loss in Clinical Decisionmaking

❖Jeng Mah, American Medical Systems, 10700 Bren Road West, Minnetonka, MN 55343, *the_mahs@comcast.net*

Key Words: Bayesian hypothesis test, posterior probability, Bayes factor, expected loss, optimal decisions

In a previous paper, I proposed a decision model to explicitly harness loss in clinical trial design (e.g., in choosing the optimal minimum detectable difference d for superiority or noninferiority). This approach requires knowledge of P(H0), or a P(H0) that can be described in terms of d, since P(H0) is otherwise meaningless in a frequentist sense. Bayesian inference relies on P(H0) and P(H1); a loss-based decision approach is therefore a natural fit in Bayesian study design. This paper considers a loss based decision model for Bayesian clinical study design. In this model, action (i.e., rejecting or accepting H0) is taken based on Bayesian inference and the losses incurred from the decision. Simple cases with parameters having conjugated priors are studied in deriving optimal decisions in hypothesis testing. Actions based on posterior probability and Bayes factor are also examined.

Constructing Bayesian Trials in a Type I Error–Controlled Regulatory World: The Good, the Bad, and the Ugly

❖Scott Berry, Berry Consultants, 3145 Chaco Canyon Drive, College Station, TX 77845, *scott@berryconsultants.com*

Key Words: Bayesian, Adaptive Trials, Medical Devices, Type I error, Priors

The use of Bayesian statistics in trial design allows for innovative adaptive designs to be done. Some strengths of the Bayesian approach include using prior information, flexibility in the adaptive sample size, flexibility in randomization probabilities and predictive probabilities. Many regulatory agencies allow and encourage Bayesian and adaptive techniques in clinical trial design, but constrain the Bayesian approach to satisfy frequentist type I error levels. In this talk I discuss the ramifications of these restrictions: some good, some bad, and some ugly!

Bayesian Designs and Therapeutic Medical Device Trials

❖Pablo Bonangelino, Food and Drug Administration, CDRH, 1350 Piccard Dr, Rockville, MD 20850, *pablo.bonangelino@fda.hhs.gov*

Key Words: Bayesian, medical device, CDRH

This talk will review the Bayesian designs which have been used in medical device trials that have been submitted to the General and Surgical Devices Branch of FDA's Center for Devices and Radiological Health (CDRH) in recent years. I will begin with an overview of the approach taken by our Branch towards Bayesian statistical methods. I will then discuss examples of Bayesian methods, which include likelihood based methods for incomplete data, Bayesian adaptive designs using predictive probability, and incorporating prior data in a confirmatory study. Regulatory considerations regarding these designs are explored. I will conclude by considering the future for Bayesian techniques in therapeutic device studies.

Examples from Medical Device Clinical Trials To Illustrate Advantages of the Bayesian Approach

❖Teresa Nelson, Technomics Research, LLC, 1815 Medina Road, Medina, MN 55356, *tnelson@technomicsresearch.com*

Key Words: Bayesian, Advantage, Medical Device, Clinical Trial, Predictive Distribution, Hierarchical Model

The presentation will include examples illustrating advantages to Bayesian methods. The examples will include: 1) Utilization of prior data to reduce overall sample size by two methods (i.e., use of an informative prior and a hierarchical model). 2) Incorporation of multiple looks via a Bayesian design allows earlier trial stopping (i.e., less overall sample) than the typical frequentist group sequential design given the same assumptions about the underlying treatment effect and non-informative prior information. 3) Utilization of the predictive distribution to calculate probability of successful trial result based on the data collected to date in the trial and assumptions about the data yet to be observed/collected in the trial. 4) Ability to combine data from multiple sources via a Bayesian Meta-Analysis (Generalized Evidence Synthesis and a Random Effects Model).

Bayesian Sample Size Calculation with Average Length Criterion Using Historical Control Data in Medical Device Trials

❖Liang Li, Genzyme Corporation, BioMedical Operation, 500 Kendall Street, Cambridge, MA 02142, *Liang.Li@genzyme.com*; Ming-Hui Chen, University of Connecticut; Hong Wang, Boston Scientific Corporation

Key Words: Bayesian, Sample Size Calculation, Average Length Criterion, Historical Data

We used the power prior and average length criterion approaches to the problem of determining the required sample size of binomial parameters for Bayesian noninferiority trial with historical control data. We demonstrate the efficacy of Bayesian-based sample-size determination method with a medical device trial example using Monte Carlo simulation.

374 Statistical Research in the VA Healthcare System ● ✪

Section on Health Policy Statistics, Section on Government Statistics, Section on Statistics in Epidemiology

Wednesday, August 1, 8:30 am–10:20 am

Determining Sample Sizes for Estimating and Comparing Item Parameters from the Rasch Model

❖Linda Young, University of Florida, Dept of Statistics, PO Box 11039, Gainesville, FL 32611-0339, *ljyoung@ufl.edu*

Key Words: power, sample size, Rasch model

In the Rasch Model, the probability of a specified response (e.g., right vs. wrong) is modeled as a function of person and item parameters. With the VA, this model is frequently used to analyze results from functional status and quality of life surveys. Yet, sample size requirements for assessing an item's difficulty or for comparing the difficulty of two items are based on only rough "rules-of-thumb" that are often neither empirically nor theoretically based. The purpose of this study is to determine the sample sizes needed to estimate a parameter with the desired precision or to compare the difficulty of two items with specified power. The sample size requirements over a range of item difficulty and distribution of person parameters are determined. Sample sizes can be more precisely determined than the "rules-of-thumb" currently prevalent in the literature.

The Impact of Imbalance in Two-level Logistic Models

❖Huanyu Chen, University of Pittsburgh, 3633 Allendale Circle, Pittsburgh, PA 15204, *huanyu.chen@va.gov*; Roslyn A. Stone, Veterans Affairs Pittsburgh Healthcare System

Key Words: random intercept (RI) model, random coefficient (RC) model, racial disparities research

We extended explicit variance formulas for fixed effects in two-level balanced binary data (Moerbeek et al. 2002, 2005) to account for imbalance both between and within clusters. Derivations were based on a linearization of 2-level RI and RC logistic models using first order marginal quasi-likelihood (MQL1). In a study of racial disparities in 30-day mortality among black and white veterans under age 65 following hospitalization for pneumonia at 149 VA hospitals (n=37,111), the site-specific proportions black ranged from 0% to 89.9% (29.1% overall). The observed between- and within-site imbalances in these VA data increased the variances of the race coefficients in the RI and RC models by factors of 1.23 and 1.16, respectively, relative to the balanced case. Compared to simulation results using second order penalized quasi-likelihood, analytic variances based on MQL1 are biased downward.

A Comparison of Methods To Impute Missing Race in VA Administrative Data

❖Maria Mor, CHERP/VA Pittsburgh Healthcare System, VA Pittsburgh Healthcare System, University Drive C (151C-U), Bldg. 28, Pittsburgh, PA 15240-1000, *maria.mor@med.va.gov*; Chunrong Cheng, University of Pittsburgh; Bevanne Bean-Mayberry, Veterans Affairs Greater Los Angeles; Gong Tang, University of Pittsburgh; Roslyn A. Stone, Veterans Affairs Pittsburgh Healthcare System

Key Words: Missing Data, Government Statistics, Multiple Imputation

VA administrative sources commonly are used to assess quality and health equity within the VA Healthcare System. However, missing racial data continue to be problematic, even after the implementation of improved data collection methods in FY2003. Using VA External Peer Review Program data for a national sample of veterans (2001–2003), we assessed the impact of alternative missing data procedures on the estimation of racial disparities in the receipt of several quality measures. Race and demographic information obtained from VA inpatient and outpatient administrative data were supplemented with race data from the Centers for Medicare and Medicaid Services. Point estimates, standard errors, and tests of significance were compared across all methods. Although qualitative changes in racial estimates were rare, the choice of statistical method affected the tests of statistical significance.

Estimating Concordance for a Hazard Scale in the Presence of Time-Varying Data

❖Peter Richardson, Baylor College of Medicine, 5213 Shady Maple Dr, Kingwood, TX 77339, *peterr@bcm.tmc.edu*

Key Words: concordance, time-varying covariates, large database

Concordance is an important summary measure of the operating characteristics of a continuous scale to measure risk of an event on the basis of subject characteristics. We examine some issues that arise in estimating the concordance between the values of a model-derived risk scale and some measure of the onset of the event in the presence of time-varying subject data. These issues include how the form of the data affect choices among equivalent alternative forms of the concordance definition as well as method of sampling observations per subject. We illustrate these issues for dichotomous event indicators (classical c-index) and for time to event (Harrell's c) in the context of data derived from hospital chart review and large healthcare system databases.

Analyzing Health Care Cost Data in the Presence of Heterogeneous Patients: Wouldn't Three Pieces Be Better Than One?

❖James F. Burgess, Jr., University of Washington, Department of Psychiatry and Behavioral Sciences, 1959 NE Pacific Street, Box 356560, Rm BB1644, Seattle, WA 98195-6560, *klsloan@u.washington.edu*; ❖Kevin L. Sloan, University of Washington, Department of Psychiatry and Behavioral Sciences, 1959 NE Pacific Street, Box 356560, Rm BB1644, Seattle, WA 98195-6560, *klsloan@u.washington.edu*

Key Words: Risk Adjustment, Sub-Classification Models, Cost Estimation

Effective prediction of health care costs for policy purposes requires balancing transparency and unbiasedness in choosing methods. Case mix information first can classify individual patients into spending type groups, then predict health care costs in each group. This is a special case of subclassification models in statistics, but computationally much simpler to implement. We use split sample validation on FY2001 VA data to compare six models to our proposed three piece OLS model using residual mean square error (RMSE), mean absolute prediction error (MAPE), and predictive ratios by deciles for evaluation across the distribution of costs. The approach we propose performs much better on these evaluation criteria. We take a simple approach to subclassification, as opposed to more complex methods previously suggested (i.e., mixture models) to achieve better cost estimation performance.

375 Risk Measures in Actuarial Science ● ✪

Section on Nonparametric Statistics, Section on Risk Analysis, Business and Economics Statistics Section

Wednesday, August 1, 8:30 am–10:20 am

Nested L-Statistics and Their Use in Comparing the Riskiness of Portfolios

❖Vytaras Brazauskas, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201-0413, *vytaras@uwm.edu*; Bruce Jones, University of Western Ontario; Madan Puri, Indiana University; Ricardas Zitikis, University of Western Ontario

Key Words: conditional tail expectation, proportional hazards transform, insurance losses, risk measure

Inspired by the problem of testing hypotheses about the equality of several risk measure values, we find that the ``nested L-statistic"—a notion introduced herein—is natural and particularly convenient. Indeed, the test statistic that we explore in this paper is a nested L-statistic. We discuss large-sample properties of the statistic, investigate its performance using a simulation study, and consider an example involving the comparison of risk measure values where the risks of interest are those associated with tornado damage in different time periods and different regions.

Risk Measures, Distortion Parameters, and their Empirical Estimation

❖Bruce Jones, University of Western Ontario, Dept of Stats and Act Sci, London, ON N6A 5B7 Canada, *jones@stats.uwo.ca*; Ricardas Zitikis, University of Western Ontario

Key Words: Premium principle, risk measure, distortion parameter, proportional hazards transform, Wang transform, conditional tail expectation

Risk measures are of considerable current interest. Among other uses, they allow an insurer to calculate a risk-loaded premium for a random loss. However, the premium principle in use by the insurer may be, at least in part, based on considerations other than risk. It is then important to quantify the degree to which the premium compensates the insurer for the risk associated with the loss. This can be done by choosing a suitable risk measure and solving for the parameter that leads to the insurer's premium. When the loss distribution is unknown, this becomes a statistical estimation problem. I will discuss the nonparametric estimation of the parameter associated with a distortion-based risk measure. It is assumed that the premium principle is known, but no information is assumed about the loss distribution, and therefore empirical estimators are used.

Weighted Premium Calculation Principles

❖ Edward Furman, University of Haifa, Hativat Givaty [illegible], Carmiel, 20100 Israel, *efurman@stat.haifa.ac.il*; Ricardas Zitikis, University of Western Ontario

Key Words: weighted transform, weighted distribution, weighted premium calculation principles

We suggest a broad class of premium calculation principles, investigate their properties and discuss statistical inferential results. The class, which we call weighted premiums, includes many well known premiums such as the net, variance, Esscher, tail conditional expectation, and tail variance premiums. The weighted premiums are closely related to the so-called weighted distributions, which have been extensively used in statistics.

Testing Hypotheses About the Equality of Several Risk Measure Values with Applications in Insurance

❖ Madan Puri, Indiana University, Department of Mathematics, Bloomington, IN 47405, *puri@indiana.edu*

Key Words: hypotheses testing, risk measure, asymptotic results

The talk explores statistical tests about the equality of risk measure values obtained using a distortion-based risk measure. We consider both the cases in which the risk measure value is specified in the null hypothesis and the case in which it is not. In the former case, one- and two-sided alternatives are considered, and in the latter case, ordered and unordered alternatives are considered. Asymptotically most powerful tests are obtained, and asymptotic distributions of the test statistics are found using results about the asymptotic distributions of the risk measure values.

Estimating Conditional Tail Expectations with Actuarial Applications in View

❖ Ricardas Zitikis, University of Western Ontario, 1151 Richmond Street North, London, ON N6B 3R7 Canada, *zitikis@stats.uwo.ca*

Key Words: Conditional tail expectation, Premium, Mean residual life, Asymptotics, Insurance, Statistical inference

We shall discuss statistical inferential tools for estimating and comparing conditional tail expectation (CTE) functions, which are of considerable interest in actuarial science. In particular, we shall discuss estimators for the CTE functions, asymptotic theory for the estimators, and then discuss constructing pointwise and simultaneous confidence intervals for the functions. Theoretical results are obtained under minimal assumptions. The general Vervaat process plays a crucial role in achieving the aforementioned goals.

376 Analysis of Binary Data and Directional Data

Biometrics Section

Wednesday, August 1, 8:30 am–10:20 am

Some New Models and Procedures for Directional Data

❖ Brett Presnell, University of Florida, Dept of Statistics, PO Box 118545, Gainesville, FL 32611-8545, *presnell@stat.ufl.edu*; Pavlina Rumcheva, C-BASS, LLC

Key Words: Directional data, Langevin distribution, von Mises distribution, Fisher distribution, Projected normal distribution, Random effect

Some new inferential procedures are presented for directional data in the general d-dimensional case, with two and three dimensional directions being of primary interest. We first consider new tests for equality of mean directions and concentrations for multisample data from the Langevin-von Mises-Fisher and the projected normal distributions. The tests for equality of mean directions in particular perform well over broad ranch of concentrations, reducing or eliminating the need to tailor the test to the underlying unknown concentration. We also discuss and fit random and mixed effects models for directional data. These models generalize the probit-normal generalized linear mixed model from the binary case of a one dimensional direction to higher dimensions.

Analysis of Binary and Continuous Outcomes in Neonatal Clinical Trials with Twin Births

❖ Michele Shaffer, The Pennsylvania State University, 600 Centerview Drive, College of Medicine, Suite 2200 A210, Hershey, PA 17033, *mshaffer@hes.hmc.psu.edu*; Kristi L. Watterberg, University of New Mexico; Allen Kunselman, The Pennsylvania State University

Key Words: Correlated data, mixed-effects models, generalized estimating equations

In neonatal trials of pre-term or low-birth-weight infants twins may represent 10-20% of the study sample. Mixed-effects models and generalized estimating equations are common approaches for handling correlated binary or continuous data. However, the operating characteristics of these methods for mixes of correlated and independent data are not well established. Simulation studies are conducted to compare mixed-effects models and generalized estimating equations to linear regression for continuous outcomes and logistic regression for binary outcomes to estimate the treatment effect in two-armed clinical trials. Additionally, the ability to estimate the magnitude of within-birth correlation for twins is examined for the correlated data methods. Data from the National Institute of Child Health & Human Development Neonatal Research Network are used for illustration.

Analysis of Relatively Large Sparse Binary Data

❖ Ko-Kang Wang, The Australian National University, Burgmann College, Canberra ACT, 2601 Australia, *kevin.wang@maths.anu.edu.au*; Susan Wilson, The Australian National University; Alan Welsh, The Australian National University; John Maindonald, The Australian National University

Key Words: Logistic regression, permutation test

Modern technological developments are enabling relatively large numbers of measurements to be made on a relatively small number of individuals. The data motivating this study were obtained to investigate interactions between known diabetes marker loci for determining thymic deletion in Non-Obese Diabetes (NOD) mice. Logistic regression analyses and permutation tests were used to investigate the genetic complexity of this trait.

Generalized Linear Mixed Models for Binary Outcome Data with a Low Number of Occurrences: Understanding When Estimation Procedures Fail

❖ Marie-Eve Beauchamp, McGill University, 1020 Pine Avenue West, Montreal, QC H3A 1A2 Canada, *marie-eve.beauchamp@mail.mcgill.ca*; Robert W. Platt, McGill University; James A. Hanley, McGill University

Key Words: Adaptive Gauss-Hermite quadrature, Penalized quasi-likelihood, Sparse binary outcome data

The limits of estimation methods for generalized linear mixed models that are available in commercial software packages have not been studied comprehensively. Penalized quasi-likelihood (PQL) is among the most commonly used methods. Some authors have recommended using adaptive Gauss-Hermite quadrature (AGHQ) as it is expected to produce less biased parameter estimates. We compared via simulations the performance of PQL and AGHQ for several settings of binary outcome data with a low number of occurrences. Estimation procedures failed for several datasets with a low number of clusters and a small cluster size. Types of failure observed are convergence problems, parameter estimates outside of their realistic range and biased parameter estimates. We attempted to characterize the circumstances in which the estimation procedures failed.

Binary Time Series Modeling with Application to Kinetic

❖Ying Hung, Georgia Institute of Technology, 1460 F willow lake dr, Atlanta, GA 30329, *yhung@isye.gatech.edu*; Chien-Fu Jeff Wu, Georgia Institute of Technology

Key Words: binary time series, goodness-of-fit

Micropipette experimentation is a new biotechnological method developed to measure the kinetic rates of cell adhesion which play an important role in tumor metastasis and cancer mutation. Traditional analysis of micropipette experiments assumes that the adhesion test cycles are independent Bernoulli trials. This assumption can often be violated in practice. In this paper, a multiple time series model incorporating random effects is developed to analyze the repeated adhesion tests. A goodness-of-fit statistic is introduced to assess the adequacy of distribution assumptions on the dependent binary data with random effects. The asymptotic distribution of the goodness-of-fit statistic is derived. Application of the proposed methodology to some real data in an T-cell micropipette experiment reveals some interesting information on the dependency between repeated adhesion tests.

Regional Spatial Cluster Detection for Repeated Measured Outcomes

❖Andrea Cook, Group Health Center for Health Studies, 1730 Minor Avenue Suite 1600, Seattle, WA 98101, *cook.aj@ghc.org*; Diane R. Gold, Channing Laboratory, Brigham and Women's Hospital, and Harvard Medical School; Yi Li, Harvard School of Public Health

Key Words: Cluster Detection, Cumulative Residuals, Repeated Measures, Spatial, Wheeze

Spatial cluster detection has become an important methodology in quantifying the effect of hazardous exposures. Previous methods have focused on outcomes that are binary or continuous. However, there are no spatial cluster detection methods proposed for repeated measured outcomes. This paper proposes a method using cumulative geographic residuals through a permutation test for repeated measured outcomes. A major advantage of this method is its ability to readily incorporate information on study participants relocation, which most cluster detection statistics cannot. Application of these methods will be illustrated by the Home Allergens and Asthma prospective cohort study analyzing the relationship between environmental exposures and repeated measured outcome, persistent wheezing in the last six months, while taking into account moving.

377 Advances in Some Classic Tests of Hypotheses ● ✪

Biometrics Section

Wednesday, August 1, 8:30 am–10:20 am

Likelihood Ratio Tests of Correlated Multivariate Samples

❖Erning Li, Texas A&M University, Department of Statistics, 3143, College Station, TX 77843-3143, *eli@stat.tamu.edu*; Johan Lim, Texas A&M University

Key Words: Correlated samples, Empirical rejection probability, Equality of mean vectors, Homogeneity of covariance matrices, Multivariate analysis, Resampling

We develop methods to compare multiple multivariate samples which may be correlated. The methods are new in the context that no assumption is made about the correlations among the samples to be compared. The samples are assumed to be multivariate normally distributed and balanced. Three types of null hypothesis are considered: equality of mean vectors, homogeneity of covariance matrices, and equality of both mean vectors and covariance matrices. Finite sample and asymptotic properties of likelihood ratio tests are derived. Following the theoretical findings, we propose a resampling procedure for the implementation of the likelihood ratio tests in which no restrictive assumption is imposed on the structures of the covariance matrices. The performance of the testing procedures is investigated using simulations and via a dental health study.

A Comparison of Node-Splitting Rules in Recursive Partitioning Analysis of Multivariate Quantitative Structure Activity Data

❖Yujun Wu, University of Medicine and Dentistry of New Jersey, 312 Hana Rd, Edison, NJ 08817, *wuy5@umdnj.edu*; Marc Genton, University of Geneva; Leonard A. Stefanski, North Carolina State University

Key Words: QSAR, classification tree, Hotelling's T2, MultiSCAM, pooled component test

Tree-structured methods have been advocated to model non-linear relationships in quantitative structure-activity relationship (QSAR) studies. One such algorithm, MultiSCAM (Keefer, 2001), was developed to analyze QSAR data with multivariate continuous responses, classifying objects into similar categories by building a tree with recursive partitioning. Hotelling's T2 test is the method used to determine splits in MultiSCAM. However, this test is not feasible when the dimension of an observation exceeds the number of observations. This problem is further exacerbated by the fact that missing values are common. We consider two alternatives, the pooled component test (PCT) proposed by Wu et al. (2006), and a simple ANOVA F test. To compare the three node-splitting tests, we introduce a comprehensive simulation design. A drug discovery data set is used to illustrate the methods.

On Identifying Minimum Efficacious Doses in Combination Drug Trials

❖Julia Soulakova, University of Nebraska-Lincoln, 340 Hardin Hall North, Department of Statistics, Lincoln, NE 68583-0963, *jsoulakova2@unlnotes.unl.edu*; Allan Sampson, University of Pittsburgh

Key Words: Average test, Closed testing, Dose response, Minimum effective dose, Step-down procedure

In the combination drug case with at least one component known to be effective at the considered doses, the FDA regulatory requirements can be satisfied by showing that the combination is superior to each component. We term such a combination efficacious. The problem is to identify the minimum efficacious combinations, which is important especially when administering higher doses of the drug can cause serious side effects. We propose a testing procedure based on the "average" test under the closed testing principle. This procedure is shown to have strong control of overall error. The simulation studies are used to assess the performance in terms of power, lack of power and family-wise error. Among non-empty minimum efficacious dose sets the quality of the procedure improves with increasing global gain average and the procedure performs the best when the set contains the highest dose.

Weighted Distance Measures for Metabolomic Data

❖ Philip Dixon, Iowa State University, 120 Snedecor Hall, Ames, IA 50011-1210, *pdixon@iastate.edu*

Key Words: cluster analysis, PCA

A typical metabolomic study measures the concentration of a very large number of cellular metabolites in a relatively small number of samples. Biologists frequently analyze metabolomic data using a cluster analysis based on Euclidian distance between samples or a Principal Components Analysis, which is indirectly based on Euclidian distance. We examine patterns of variation between replicate samples in two data sets and find that the s.d. is approximately proportional to the mean. Based on this, we develop three weighted measures of distance that better differentiate treatments. The Canberra distance is equivalent to one of these weighted measures. A modification of the weights provides a reasonable analysis of samples with below detection-limit values.

Nondeletion Approach to Detect Discordant Subjects in Repeated Measurements

❖ Jungwon Mun, University of Wisconsin-Madison, Department of Statistics, 1300 University ave, Madison, WI 53706, *mun@stat.wisc.edu*

Key Words: diagnostics, repeated measurements, mixed effect model, discordant subject

For functional data, there are few diagnostic tools. Some methods have been suggested which adjust current methods for regression data to functional data. However, diagnostics previously proposed for functional data do not generalize well to random effects and nonlinear models. This paper introduces a diagnostic tool which indicates the source of deviation in subjects and works for all types of repeated measurement data with linear or nonlinear mixed models. The new method utilizes the residual sum of squares at each level of nesting from mixed effect models and explores the data both subject-wise and observation-wise. Examples from Orthodontic growth and biomedical engineering are presented.

A Combined Test for the Equality of Means

❖ Yvonne Zubovic, Indiana University Purdue University Fort Wayne, 2101 Coliseum Blvd East, Fort Wayne, IN 46805-1445, *zubovic@ipfw.edu*; Chand Chauhan, Indiana University Purdue University Fort Wayne

Key Words: combined tests, mean comparisons

In many situations it may be desirable to combine the results of several independent tests for a hypothesis of interest. The reason for combining tests may be that some data are lost or not conveniently available, or that the individual experiments are too dissimilar to combine. The combined test may be a function of the attained significance levels, say P1, P2,., Pk, of k independent tests or a function of their corresponding test statistics T1, T2, ., Tk. Combining the attained significance levels leads to a nonparametric test based on the Chi Square distribution. However, a test based on a function of the test statistics requires knowledge of the distribution of each Ti, as well as the distribution of the function of the statistics. In the present paper the authors discuss the above mentioned techniques of combining tests when testing the equality of means.

Analysis of One-Way Layout of Overdispersed Count Data

❖ Krishna Saha, Central Connecticut State University, Department of Mathematical Sciences, 1615 Stanley Street, New Britain, CT 06050, *sahakrk@ccsu.edu*

Key Words: Count data, Double extended quasi-likelihood, Extended quasi-likelihood, Homogeneity of the means, Overdispersion

This paper develops procedures for testing the homogeneity of the means of several groups of count data with overdispersion. The score type tests based on the models that are specified by only the first two moments of the counts are obtained using quasi-likelihood (QL), extended quasi-likelihood (EQL), and double extended quasi-likelihood (DEQL). Simulations are then used to study the comparative behavior of these statistics compared to the statistic based on the negative binomial model, in terms of size, power and robustness for departures from the data. These simulations demonstrate that the statistic based on the DEQL holds the nominal level well, and it shows some edge in power over the other statistics, and, in particular, it performs much better than the commonly used statistic based on the QL. Finally, applications to toxicological and biological data are given.

378 Bayesian Approaches in Drug Development ● ✪

Biopharmaceutical Section

Wednesday, August 1, 8:30 am–10:20 am

A Revisit to Sample Size Computation

❖ Ruoyong Yang, Pfizer Inc., 235 E 42nd St, New York, NY 10017, *ruoyong.yang@pfizer.com*

Key Words: sample size computation, non-informative prior, average power, simulation, conditional power, sample size re-estimation

In phase III clinical trials, we regularly compute sample size based on previously observed treatment effect. In this paper, sample size computation is explored with consideration of uncertainties around estimated treatment effects and corresponding standard deviations. Bayesian framework is introduced to compute average power under standard non-informative priors. The numerical integrations are easily implemented with a simple SAS macro. Standard sample size computation is found over optimistic in most cases. Sample size required under average power framework is explored. Standard conditional power computation is also found over optimistic in most cases. Sample size re-estimation method is incorporated into the average power framework. Substantially different additional sample size is required in some cases, compared with standard sample size re-estimation method.

Joint Exploration of Haplotype Information and High-Order SNP-Phenotype Associations via Gibbs Sampling

❖ Jing Zhao, Merck & Co., Inc., UG1C46 PO Box 1000, North Wales, PA 19454-1099, *jzhao@alumni.upenn.edu*

Key Words: Bayesian Variable Selection, Hierarchical Models, Gibbs Sampling, Metropolis-Hasting Algorithm, Single Nucleotide Polymorphism (SNP)

Haplotypes are believed to have important implications for identifying complex disease. However, obtaining haplotypes experimentally is considerable expensive. This paper presents a hierarchical Bayesian model and the associated variable (genotype) selection procedure that simultaneously reconstruct the haplotypes and locate the multi-locus genetic contributors to complex disease. We demonstrate that this approach is well suited to large model spaces and has reasonable power to detect high order genotype-phenotype associations. Gibbs sampling is implemented and the model with the highest visiting frequency is selected as the best model.

Using the Bayesian Approach for Pharmaceutical Process Optimization

❖ Aili Cheng, GlaxoSmithKline, 709 Swedeland Road, UE0433, King of Prussia, PA 19406, *aili_2_cheng@gsk.com*; John Peterson, GlaxoSmithKline; Younan Chen, GlaxoSmithKline; Gregory Stockdale, GlaxoSmithKline

Key Words: Response surface optimization, Bayesian analysis, Quality by design, Design Space

FDA and ICH recently published several documents that encourage the drug process understanding "from pharmaceutical development studies and manufacturing experience" to "support the establishment of the design space, specifications, and manufacturing control". It is specially emphasized that "quality should be built by design" and risk management, the key component of which is risk assessment. In this presentation, the Bayesian posterior predictive approach developed by Peterson (2004) is utilized to identify design space and a reliable operating region for the production of a pharmaceutical drug substance or their intermediates. This approach not only provides a risk assessment but also accounts for modeling uncertainty and the correlation structure of the multiple responses, which can not be done by the traditional, frequentist approaches.

Bayesian Spines To Infer Population Mean and Tolerance Region from Multivariate Longitudinal Data

Qianqiu Li, Wyeth; ❖Yonggang Zhao, Wyeth, 500 Arcola Rd, A4034, Collegeville, PA 19426, *zhaoy2@wyeth.com*

Key Words: bayesian, confidence region, tolerance region, MCMC, longitudinal data, splines

Spine models are widely used for the analysis of longitudinal data. In this talk, we propose a high-dimensional spline model in mixed model framework to address two common issues arising in clinical safety studies: population mean inference and tolerance region construction for screening of individuals with unusual responses, based on multivariate longitudinal data. Bayesian analysis through MCMC samplers is employed for estimation and inference. Using multivariate central limit theorem and Monte Carlo technique, we obtain approximate confidence and tolerance regions through time.

Bayesian Hierarchical Models for Detecting Safety Signals in Clinical Trials

❖Hong A. Xia, Amgen Inc., One Amgen Ctr Dr, MS 242A, Thousand Oaks, CA 91320, *hxia@amgen.com*; Haijun Ma, Amgen Inc.

Key Words: Bayesian Hierarchical Models, Clinical Trials, Drug Safety, Multiplicity, Signal Detection

Detection of safety signals from routinely collected adverse event data in clinical trials is a critical task in drug development. How to deal with the multiplicity issue in such a setting is a challenging statistical problem. Without multiplicity considerations, there is a potential for an excess of false positive signals. On the other hand, traditional ways of adjusting for multiplicity may fail to flag important signals too often. Bayesian hierarchical modeling brings some promise. It allows for explicitly modeling AEs with the existing coding structure so that they can borrow strength from each other depending on the actual data. Following the work by Berry and Berry (2004), we implement the Bayesian hierarchical mixed model for subject incidence and extend this model for exposure adjusted incidence or exposure adjusted event rate, using drug exposure time between patients.

Controlled Bayesian Optimal Designs for Dose-Response Studies

❖Wei Zhu, State University of New York at Stony Brook, Dept of Applied Math and Statistics, Stony Brook, NY 11794-3600, *zhu@ams.sunysb.edu*; Jiaqiao Hu, State University of New York at Stony Brook; Weng-Kee Wong, University of California, Los Angeles

Key Words: Bayesian multiple-objective optimal design, dose response study, logit model

Dose-response studies are routinely conducted to determine viable dose levels for newly developed therapeutic drugs. Due to safety, efficacy, and experimental design considerations, practical constraints are often imposed on (1) dose range, (2) dose levels, (3) dose numbers, (4) dose proportions, and (5) potential missing observations. We propose the controlled Bayesian multiple-objective optimal designs satisfying one or more of these practical constraints, for dose response studies. The resulting controlled optimal designs satisfying these realistic constraints can be readily adopted by the pharmaceutical researchers for optimal estimation of the parameters of interest such as the median effective dose level or the threshold dose level. We demonstrate our results and methodology through the logit dose response model.

A Bayesian Approach to Utilizing Prior Data in New Drug Development

❖Todd Coffey, Amylin Pharmaceuticals, Inc., 9360 Towne Centre Drive, San Diego, CA 92121, *todd.coffey@amylin.com*; Larry Z. Shen, Amylin Pharmaceuticals, Inc.; Wei Deng, Amylin Pharmaceuticals, Inc.

Key Words: clinical trials, event rate, patient-year observation, Poisson distribution, relative risk, safety analysis

We propose a Bayesian method to combine safety data collected from two drug development programs using the same active drug substance but for different indications, formulations, or patient populations. The key concept of our method is to use data from the previous program to construct a posterior distribution that will in turn serve as a prior distribution for the new program. This updated prior down weights data from the previous program to emphasize the new program. We have tested this approach using data from a Phase 2 study that was conducted for a new indication of an approved drug. The results indicate that the estimated safety risk level was affected both by the observed event rates and the extent of exposure. This approach appropriately characterizes the safety profile across the two development programs and properly contextualizes new safety signals from the new program.

379 Linear Models in Clinical Trials

Biopharmaceutical Section

Wednesday, August 1, 8:30 am–10:20 am

Inverse Prediction: A Clinical Application

❖Jayawant Mandrekar, Mayo Clinic, 2665 Boulder Ridge Drive NW, Rochester, MN 55901, *mandrekar.jay@mayo.edu*

Key Words: Inverse Prediction, Regression

An important feature of regression methodology is in the area of prediction. Investigators are typically interested in predicting a value of a response variable (Y) based on the known value of the predictor variable (X). However, sometimes there is a need to predict a value of the predictor variable (X) based on the known value of the response variable (Y). In such situations, it is improper to simply switch the roles of the response and predictor variables to get the desired predictions i.e., regress X on Y. This is because the primary assumption that X is measured without error and Y is a dependent, random and normally distributed variable is violated. Inverse prediction is a method that accounts for the underlying assumptions while estimating or predicting X from a known Y. This approach including calculations for the 95% confidence limits will be illustrated using clinical data.

An Approach to Obtaining Initial Values for the Covariance Parameters for Repeated Measures Analysis with Unstructured Covariance Model

❖Kaifeng Lu, Merck & Co., Inc., 217 Vasser Dr, Piscataway, 08854, *kaifeng_lu@merck.com*

Key Words: repeated measures, covariance structure, missing at random

In randomized clinical trials, patient data are usually collected over time. Dropouts and the resulting missing data often pose a challenge in data analysis. Repeated measures model can provide valid inference when the missing data are missing at random. An unstructured covariance model is often used to avoid bias due to misspecification of the covariance structure. Convergence problems may arise for various reasons. In some situations, convergence may be achieved via the use of reasonable initial values for the covariance parameters. One approach to obtaining initial values is to fit a series of univariate analyses of covariance models for the measurement collected at the current time point conditional on baseline covariates and measurements collected at previous time points. The covariance matrix for the repeated measures can then be reconstructed from the sequence of ANCOVA models.

Permutation Test-Based Significance Levels for Repeated Measures

❖Katie Spencer, Zimmer, Inc., PO Box 508, Warsaw, IN 46581-0708, *katie.spencer@zimmer.com*; Daniel Reyner, Zimmer, Inc

Key Words: repeated measures, distribution free, permutation test, model assumptions, clinical trial, 2 group comparison

Repeated measures analyses are common to clinical trials which seek regulatory approval. However, violation of model assumptions related to the distribution of residual error may require abandonment of the statistical model. This paper provides an example of the use of permutation tests to generate measures of statistical significance in repeated measures analyses. This method allows for a prespecification of the statistical model which will be used in final analysis, achieving estimates of statistical significance which are distribution-free. The example focuses on comparisons of repeated measures between two groups.

A Multiplicative Heteroscedasticity-Based Transformation To Reduce Heteroscedasticity in General Linear Models

❖Mitchell Rosen, Omnicare Clinical Research, 630 Allendale Rd., King of Prussia, PA 19406, *mitch.rosen@omnicarecr.com*

Key Words: transformation, Box-Cox, multiplicative heteroscedasticity, homogeneity of variance, delta method, general linear model

Data analysts commonly use transformations to achieve homogeneity of variance, linearity, additivity, and/or normality. Box and Cox (1960) proposed estimating a single parameter power transformation model using maximum likelihood methods. Their method is optimum when normality of residuals is desired. In this paper the power transformation model is estimated using Harvey's (1976) multiplicative heteroscedasticity model. This method is optimum when homoscedasticity is desired. Estimation is accomplished using generalized linear model methods. Performance of the transformation vis-‡-vis Box-Cox is assessed using Monte Carlo methods.

Choosing an Appropriate Variance-Covariance Structure for Small Sample Size Data

❖Zhaoling Meng, sanofi-aventis, 181 Buttonwood Dr, Piscataway, NJ 08854, *zhaoling.meng@sanofi-aventis.com*; Xun Chen, sanofi-aventis; Robert Kringle, sanofi-aventis

Key Words: variance-covariance structure selection, repeated measurements, type I error inflation, loss of power, small sample size

In repeated measurement analyses, variance-covariance structure is considered having potential impact on modeling main interests, estimating/testing fixed effects. Choosing appropriately is no easy task. Relying on information criteria such as AIC/AICC/BIC is a common practice but with known limitations, such as overfitting/underfitting. Type I error inflation is often observed. Or a pre-assumed general variance-covariance structure is used, such as un-structured, without consulting data to avoid selection. Such over-generality could lead to significant loss of power, especially for studies with relatively small sample sizes. Simulations were conducted for typical phase I settings, impacts of above two common practices were explored with attention in balancing type I error inflation and loss of power. Benefit of heterogeneity among treatments if pre-knowledge suggests was investigated.

Modeling Covariance Structure in Unbalanced Longitudinal Data

❖Min Chen, Texas A&M University, Statistics Department, 1100 Hensel Dr., Apt. X2H, College Station, TX 77840, *mchen@stat.tamu.edu*; Jianhua Z. Huang, Texas A&M University

Key Words: logitudinal, unbalanced, covariance, reparameterization, EM algorithm, decomposition

Modeling covariance structure is important in efficient estimation for longitudinal data models. Pourahmadi (1999,2000) promoted to use modified Cholesky decomposition as an unconstrained reparameterization of the covariance matrix. The new parameters have transparent statistical interpretations and are easily modeled using covariates. However, this approach is not directly applicable when the longitudinal data are unbalanced, because a Cholesky factorization for observed data coherent across all subjects usually does not exist. We overcome the difficulty by treating the problem as a missing data problem and employing a generalized EM algorithm to compute the ML estimators. We illustrate our method by reanalyzing Kenward's (1987) cattle data and conducting a simulation study.

Permuted Block Randomization and Balance in Clinical Trials

❖John (Joong Hee) Han, Centocor, Inc., 200 Great Valley Parkway, Malvern, PA 19355, *jhan9@cntus.jnj.com*; Lee-Lian Kim, Centocor, Inc.

Key Words: block randomization, balance, clinical trial

There are different methods to randomly allocate subjects to treatment arms in a clinical trial. One method that is commonly used is permuted block randomization. The major advantage of permuted block randomization is that when all blocks are completed at the end of a clinical trial, there is complete balance between treatment groups. This is true for the whole trial as well as for any stratification factor that may have been incorporated into the randomization scheme. However, if there are few subjects in stratification cells, incomplete blocks could exist, which may lead to imbalance among treatment groups. This presentation will explore the degree of imbalance that may arise and if this imbalance can affect the results of a trial.

380 Applications in Statistical Genetics ●

ENAR, Section on Teaching Statistics in the Health Sciences, Biometrics Section

Wednesday, August 1, 8:30 am–10:20 am

Gene-Environment Interaction Test for Dichotomous Traits in Nuclear Families

❖ Thomas Hoffmann, Harvard University, 655 Huntington Ave, Dept of Biostatistics, Boston, MA 02115, *tjhoffm@fas.harvard.edu*; Nan Laird, Harvard University

Key Words: Gene environment interaction, Genetics, Sib pairs

We first propose an extension of the Lake and Laird permutation test for trios to sibpairs, and develop an alternative empirical variance. Extending to discordant sibs leads to an invalid test statistic, but sibpairs with at least one affected is valid. We explore the degree of bias in this test to determine when it will be negligible. Then, due to limitations in the test, we explore the use of the Vansteelandt and Lange gene-environment interaction test on dichotomous traits. We compare these results to other gene-environment interaction tests in trios or sibpairs.

Oracle and Adaptive Compound Decision Rules for False Discovery Rate Control

❖ Wenguang Sun, University of Pennsylvania, 503 Blockley Hall 423 Guardian Drive, Department of Biostatistics, Philadelphia, PA 19104, *wsun@cceb.med.upenn.edu*; Tony Cai, University of Pennsylvania

Key Words: Adaptive procedure, compound decision rule, false discovery rate

We develop a compound decision theory framework for multiple testing problems and derive an oracle rule based on the z-values that minimizes the false non-discovery rate (FNR) subject to a constraint on the false discovery rate (FDR). It is shown that many commonly used multiple testing procedures, which are p-value based, are inefficient. An adaptive procedure based on the z-values is proposed. It is shown that the z-value based adaptive procedure asymptotically attains the performance of the z-value oracle procedure and is more efficient than the conventional p-value based methods. Numerical performance of the adaptive procedure is investigated using both simulated and real data. In particular our method is demonstrated in an analysis of the microarray data from a HIV study that involves testing a large number of hypotheses simultaneously.

A General Tree-Based Regression Model for Multivariate Outcomes

❖ Kai Yu, National Institutes of Health, DCEG,NCI, Biostatistics Branch, 6120 Executive Blvd, Executive Plaza South, Room 8050, Rockville, MD 20852-7244, *yuka@mail.nih.gov*; William Wheeler, Information Management Service Inc; Qizhai Li, National Institutes of Health; Jinbo Chen, University of Pennsylvania

Key Words: tree model, genetic association study, multivariate outcome

In genetic association studies of complex diseases, we usually have multiple phenotypic measurements for the outcome. Those measurements could be correlated and share a common genetic basis. Therefore, it might be more powerful to analyze multivariate phenotypes jointly rather than to study each individual phenotype separately. Here we develop a tree-based regression model for that purpose. The new model builds on the framework of a recently developed partially linear tree-based regression model and extends it to multivariate outcomes. Simulation studies are conducted to evaluate the type I error rate and the power of the proposed model. We also demonstrate its application through a real example.

Using a Hybrid-Modified Genetic Algorithm To Find Feasible Regions of a Desirability Function

❖ Wen Wan, Virginia Polytechnic Institute and State University, 1835 Grayland Street, Apt. 6, Blacksburg, VA 24060, *wenw@vt.edu*; Jeffrey B. Birch, Virginia Polytechnic Institute and State University

Key Words: Desirability Function, Genetic Algorithm, Hybrid Modified Genetic Algorithm, Multi-response Optimization (MRO), Nelder-Mead Simplex Algorithm, Response Surface Methodology (RSM)

The multiresponse optimization (MRO) problem is quite common in real applications. Most of the MRO techniques such as the desirability function method by Derringer and Suich (1980) are utilized to find optimal solutions or near-optimal solutions by using some appropriate optimization routine such as the Nelder-Mead simplex algorithm with different starting points. However, practitioners usually prefer to identify all of the near-optimal solutions, or all feasible regions, because some feasible regions may be more desirable than others based on practical considerations (Myers et al., 2004). In this paper, we demonstrate the use of a hybrid modified genetic algorithm to generate all feasible regions for the desirability function without limitation to the number of factors. A case study illustrates that our procedure successfully defines all feasible regions.

A Study of Variability in Mass Spectrometry Data

❖ Ann L. Oberg, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, *oberg.ann@mayo.edu*; Douglas W. Mahoney, Mayo Clinic; Jeanette E. Eckel-Passow, Mayo Clinic; Terry M. Therneau, Mayo Clinic; Christopher J. Mason, Mayo Clinic; Kenneth L. Johnson, Mayo Clinic; K. Sreekumaran Nair, Mayo Clinic; H. Robert Bergen, Mayo Clinic

Key Words: proteomics, mass spectrometry, variability, isotope labeling

Characterization of the human proteome is a resource of tremendous potential to biological research. Shotgun proteomics via mass spectrometry is a powerful technology for study of the proteome; it has the potential to lead to a non-invasive screening mechanism of proteins in easily accessible body fluids. To effectively use statistical methodologies, it is important to understand and characterize the variance structure of the resulting data. We present findings from experiments performed via stable isotope labeled 16O/18O and iTRAQ approaches. Linear models were used to model systematic effects and evaluate variability at various levels of data acquisition. Relationships with respect to properties of the sample were evaluated as well as proteome coverage. We found variability to be related to abundance and, not surprisingly, to the number of data points available for quantification.

An Approach to Infer Gene Interaction Networks

❖ Alina Andrei, University of Wisconsin-Madison, 1300 University Avenue, Biostatistics and Medical Informatics, Madison, WI 53706, *aandrei@biostat.wisc.edu*; Christina Kendziorski, University of Wisconsin-Madison

Key Words: gene association networks, gene interactions, gaussian graphical models, partial correlation

Complex traits ranging from mRNA expression to disease phenotype are affected by multiple genes, and importantly, by gene interactions (GIs). Gene association networks (GANs) provide a simple, yet effective, summary of relationships among genes. In GANs, nodes are genes and edges represent association often measured by correlation or, as in Gaussian graphical models (GGMs), conditional dependence given all other nodes within a multivariate normal distribution. Although useful, GANs are limited since the association measures quantify a linear relationship among nodes. We propose an approach, applicable for n< p, that uses the correlation and partial correlation matrices to extend the class of GGMs and accommodate GIs. Comparisons between the resulting GI networks and traditional GGMs illustrate the advantages of the proposed approach.

Cluster-Based Phenotype Construction for Linkage Heterogeneity

❖Cavan Reilly, The University of Minnesota, A460 Mayo Building MMC 303, 420 Delaware St SE, Minneapolis, MN 55455, *cavanr@biostat.umn.edu*

Key Words: linkage analysis, locus heterogeneity

We describe a method for defining phenotypes for complex diseases that may exhibit locus heterogeneity and we use the method to identify a locus related to asthma susceptibility. Unlike other approaches to locus discovery in the presence of heterogeneity, this method seeks loci that segregate in all or most ascertained families while recognizing that other genes and environmental factors that modify the action of the common gene may vary across families. The method is based on seeking groups of families that differ, between groups, in the way affected individuals express the genotype. Then we use the distance of each individual to the cluster center for his family to define a quantitative trait. This quantitative trait is then subjected to a genome scan using variance components methods. The method is applied to a dataset of 27 families and a novel locus was identified.

381 Bayesian Methods in the Social Sciences ● ✪

Section on Bayesian Statistical Science

Wednesday, August 1, 8:30 am–10:20 am

A Bayesian Method for IRT Model Selection

❖Sherwin Toribio, University of Wisconsin-La Crosse, Mathematics Department, 1725 State St., La Crosse, WI 54601, *toribio.sher@uwlax.edu*

Key Words: IRT models, Bayesian methods, Bayes Factor, Hierarchical model, Gibbs Sampling

Item Response Theory (IRT) models are commonly used in educational and psychological testing. These models are mainly used to assess the (latent) abilities of examinees and the effectiveness of the test items in measuring this underlying trait. Several models are currently available for this purpose. In this paper, a Bayesian method to compare different IRT models will be discussed. In particular, the Bayes factor will be employed to determine which model is more appropriate for the data.

A New Bayesian Approach for Misidentification with Applications for Survey Data

Simon Cheng, University of Connecticut; ❖Yingmei Xi, University of Connecticut, 215 Glenbrook Road, Unit 4120, Storrs, CT 06269, *yingmei@stat.uconn.edu*; Ming-Hui Chen, University of Connecticut

Key Words: Gibbs sampling, Latent class model, Posterior distribution

Social scientists often rely on survey data to examine group differences in various outcomes. An inherent problem in survey data, however, is the potential misidentification of group membership due to poorly trained interviewers, inconsistent responses, or simple errors in marking questions. In data that contain drastically unequal sub-sample sizes, the detrimental consequences of misidentification can be considerable, especially concerning groups with a small sample size. In this study, we develop a new mixture model that allows researchers to address the problem using the data that they have in hand. By supplying additional information from the data, this two-stage model is estimated using a Bayesian method. As anticipated, the more information we supply to adjust for group membership, the better the model performs. We conduct sensitivity analysis on choices of priors.

Model-Based Clustering of non-Gaussian Longitudinal Data

❖Miguel Juarez, University of Warwick, Department of Statistics, Coventry, CV4 7AL United Kingdom, *m.a.juarez@warwick.ac.uk*; Mark F.J. Steel, University of Warwick

Key Words: autoregressive modelling, employment growth, hierarchical prior, model comparison, skewness

In this paper we propose a model-based method to cluster units within a panel. The underlying model is autoregressive and non-Gaussian, allowing for both skewness and fat tails, and the units are clustered according to their dynamic behavior and equilibrium level. Inference is addressed from a Bayesian perspective and model comparison is conducted using the formal tool of Bayes factors. Particular attention is paid to prior elicitation and posterior propriety. We suggest priors that require little subjective input from the user and possess hierarchical structures that enhance the robustness of the inference. Two examples illustrate the methodology: one analyses economic growth of OECD countries and the second one investigates employment growth of Spanish manufacturing firms.

Markov Chain Monte Carlo–Based Mixture Modeling Applications in Psychometrics

❖Dipendra Subedi, Michigan State University, 1422 Spartan Village, Apt E, East Lansing, MI 48823, *subedidi@msu.edu*

Key Words: Bayesian Methods, Mixture Modeling, Psychometrics, MCMC

There has been growing interest in the modeling of complex data using Bayesian methods over the past few years. Psychometric applications comprise many situations in which data come from two or more distinct populations. Bayesian mixture modeling is a powerful technique to model such data using a mixture of distributions. Although Markov Chain Monte Carlo based mixture modeling has been used extensively in other disciplines (e.g., Biostatistics, Engineering), it has received limited attention in social science. This paper reviews prior work in Bayesian mixture modeling and documents the challenges and potential solutions. With the examples from psychometrics, this study illustrates the use of WinBUGS in such modeling. This study has a potential to disseminate Bayesian mixture modeling approach in psychometrics and facilitates the modeling of complex latent variables of natural phenomenon.

Bayesian Analysis of Repeated Behaviors Using Event-Dependent Frailty Models

❖Getachew Dagne, University of South Florida, 13201 Bruce B Down Blvd, Tampa, FL 33612, *gdagne@hsc.usf.edu*

Key Words: Event dependent, emotion regulation, random effects, Weibull model, Bayesian inference, social interaction

In social behavioral studies, we commonly encounter repeated displays of behaviors (events) data. Statistical methods for the analysis of such data use either parametric or semi-nonparametric proportional hazards model, modified to include frailty which summarizes the dependence of recurrent times within a dyad. These random frailty effects typically model heterogeneity among dyads. However, dyad-specific random frailty may not capture event dependence that occurs because of ordering of event occurrences. The occurrence of one behavior (event) may make further occurrence of the behavior more or less likely based on the characteristics of dyads. This paper presents an event-dependent frailty model for recurrent sojourn time data in the Bayesian context. Our approach is illustrated by a dataset relating to families with children who show emotional regulation problems.

A Joint Multivariate Ordered Probit and Tobit Modeling of Households' Vehicle Choice and Usage

❖Hao Fang, University of California, Irvine, Department of Economics, 3151 Social Science Plaza, Irvine, CA 92697-5100, *hfang@uci.edu*

Key Words: multivariate ordered probit, multivariate tobit, discrete/continuous, vehicle choice, residential density

The paper quantifies how residential density influences households' vehicle fuel economy choices and vehicle usage. Traditional nested logit modeling of vehicle choice becomes cumbersome with large numbers of vehicles and vehicle categories. I propose a more flexible method of modeling vehicle holdings in terms of number of vehicles in each category, with a Bayesian multivariate ordinal response system. I also combine the ordered equations with multivariate tobit equations to jointly estimate vehicle type/usage demand in a reduced form, offering a simpler alternative to the traditional discrete-continuous analysis. The estimation results show that the impact of density on households' vehicle choice is small but statistically significant, and comparable to income effect. The method used to estimate the system can also be applied to other discrete-continuous problems.

Fusion, Pro and Con: Model-Based Fusion of MaxDiff and Ratings Data to Put Respondent Choice Parameters on a Common Scale and Consistency with Random Utility Theory

❖Peter Lenk, The University of Michigan, Ross Business School, Ann Arbor, MI 48109, *plenk@umich.edu*; Lynd Bacon, Lynd Bacon Associates

Key Words: Bayesian Inference, Random Utility Theory, Psychometric Scaling, Ordinal Data, Marketing, Missing Data

MaxDiff scaling, sometimes called Best/Worst analysis, is a special type of choice-based conjoint analysis that involves having respondents pick "best" and "worst" items from choice sets produced from an experimental plan. A weakness of MaxDiff as it is typically applied is that scale location differences between subjects are lost. This shortcoming can be assuaged by augmenting the choice data with information about scale origin, like ratings data. There are different ways of doing this, but not all are consistent with Random Utility Theory (RUT), a cornerstone of preference modeling. In our talk we build on our 2006 JSM results to describe specifications consistent with and not consistent with RUT, compare results from applying them to those obtained from conventional procedures, and indicate the conditions under which the different specifications are to be preferred.

382 Nonparametric Inference for Correlated Data ✪

Section on Nonparametric Statistics

Wednesday, August 1, 8:30 am–10:20 am

Spectral-Based Tests for Periodicities

❖Lai Wei, The Ohio State University, 1958 Neil Avenue, Cockins Hall Room 404, Columbus, OH 43210, *weil@stat.ohio-state.edu*; Peter F. Craigmile, The Ohio State University

Key Words: Global and local tests, harmonic process, periodogram, tapering, smoothing splines, power calculations

We investigate tests for periodicity based on a spectral analysis of a time series. We contrast between global and local spectral-based tests. A global test uses the spectral estimates at all sample points, whereas a local test uses a subset of the sample points available. Standard global tests for periodicity are often based on the assumption of a Gaussian IID error process. Using a smoothing spline approach, we extend the global test to the non-IID case. We compare this approach to a number of local tests for periodicity such as the local F test, a test commonly used in hearing sciences, and Thomson's multitaper F test. Using regression-based F tests, we demonstrate that asymptotic size and power calculations can be made for some of these tests. For all the tests, we also compare the size and power at finite sample sizes, under a number of different experimental conditions.

A Nonparametric Time Reversibility Test Based on U-Statistics

❖Nusrat Jahan, James Madison University, Department of Mathematics and Statistics, MSC 1911, 319 Roop Hall, Harrisonburg, VA 22807, *jahannx@jmu.edu*; Yuji Tomita, James Madison University

Key Words: Time reversibility, U-statistics, Distance-based statistic, nonparametric

A new time series reversibility test is proposed based on a distance measure between the joint density function of the series and the series with reversed time indices. The distance measure is estimated unbiasedly and consistently within the framework of U-statistics, which can be shown to have an asymptotic normal distribution. The test does not require existence of moments higher than second order. Simulation results compare the power of the proposed test with the existing time reversibility tests.

Asymptotic Properties of Discrete Fourier Transforms of Spatial Processes

❖Soutir Bandyopadhyay, Texas A&M University, 204 Church Avenue, apt 3, College Station, TX 77840, *soutir@stat.tamu.edu*; Soumendra N. Lahiri, Texas A&M University

Key Words: Asymptotics, DFT, Spatial Processes

We investigate the joint asymptotic distribution for the Discrete Fourier Transformations (DFT) of spatial data under pure and mixed increasing-domain spatial asymptotic structure. Here the design points are specified by certain stochastic spatial designs driven by sequence of iid random vectors. We show that the asymptotic distribution of the DFT is Gaussian with the asymptotic variances depending on the design density. We also investigate the asymptotic independence of the DFTs of the spatial data with unequal frequencies.

Adaptive Density Estimation Under Dependence

❖Olivier Wintenberger, SAMOS, Centre Pierre Mends France, 90 Rue de Tolbiac, Paris Cedex 13, 75634 France, *olivier.wintenberger@univ-paris1.fr*; Irene Gannaz, University Joseph Fourier

Key Words: Adaptive estimation, Asymptotic minimax, Hard thresholding, Wavelets, Weak dependence, Nonparametric density estimation

Assume that \$(X_t)_{t\in\Z}\$ is a real valued time series admitting a common marginal density \$f\$ with respect to Lebesgue measure. Donoho {\it et al.}~(1996) propose a near-minimax method based on thresholding wavelets to estimate \$f\$ on a compact set in an independent and identically distributed setting. The aim of the present work is to extend this methodology to different weakly dependent cases. Bernstein's type inequalities are proved to be sufficient to extend near-minimax results. Assumptions are detailed for dynamical systems and under the \$\eta\$-weak dependence condition from Doukhan & Louhichi (1999). The threshold levels in our estimator integrates the dependence structure of the sequence \$(X_t)_{t\in\Z}\$ through one parameter \$\gamma\$. The near minimaxity is obtained for \$\L^p\$-convergence rates (\$p\ge 1\$). An estimator of \$\gamma\$ is obtained by cross-validation procedure.

Model Selection for Multivariate Smoothing Splines with Correlated Random Errors

❖Eren Demirhan, North Carolina State University, 1001 Avent Hill, APT. B7, Raleigh, NC 27606, *edemirh@ncsu.edu*; Hao Zhang, North Carolina State University

Key Words: COSSO, SS-ANOVA, Variable Selection, Mixed Models, Correlated Data

Model selection in nonparametric regression is a difficult problem, which becomes more challenging for correlated data such as longitudinal data and repeated measurements. In the framework of smoothing spline analysis of variance, we propose a unified approach to simultaneously selecting variables and estimating model parameters and covariance structures. The new method, as a generalization of the component selection and smoothing operator (Lin and Zhang 2006), imposes a soft-thresholding penalty on functional components for sparse estimation and take into account covariance structure at the same time. We propose an efficient algorithm in the framework of mixed effects which can be implemented by any standard software package. In particular, extensive work is done on the selection of tuning parameters. The performance of the new method is demonstrated through simulations and examples.

An Entropy-Derived Measurement of Dependence Between Two Groups of Random Variables

❖ Nan Kong, Educational Testing Services, 270 Hampshire Dr, Plainsboro, NJ 08536, *nkong@ets.org*

Key Words: Entropy, Conditional entropy, Dependence coefficient, Partial Dependence coefficient, Semi-partial Dependence coefficient

Linear relationship among random variables has been fully explored in the past. This paper looks into dependence of one group of random variables upon another group of random variables using (conditional) entropy. A new measurement of range between zero and one, called K-dependence coefficient or dependence coefficient, is defined using (conditional) entropy. This paper shows that the dependence measured by K-dependence coefficient includes both linear and non-linear dependence between two groups of random variables. I.e. K-dependence coefficient totally measures the dependence, not part of the dependence, between the two groups of random variables. Furthermore, the concept of K-dependence coefficient is extended by defining Partial K-dependence coefficient and semi-partial K-dependence coefficient. Properties of partial and semi-partial K-dependence coefficients are also explored.

One-Sample Location Tests for Multilevel Data

❖ Denis Larocque, HEC Montreal, 3000 ch Cote Sainte Catherine, Montreal, H3T 2A7 Canada, *denis.larocque@hec.ca*; Jaakko Nevalainen, University of Tampere; Hannu Oja, University of Tampere

Key Words: Multilevel data, Hierarchical data, Sign test, Wilcoxon signed-rank test, Clustered data, Asymptotic efficiency

In this paper, we consider testing the location parameter with multilevel (or hierarchical) data. A general family of weighted test statistics are introduced. This family includes extensions to the case of multilevel data of familiar procedures like the t, the sign and the Wilcoxon signed-rank tests. Under mild assumptions, the test statistics have a null limiting normal distribution which facilitates their use. An investigation of the relative merits of selected members of the family of tests is achieved theoretically by deriving their relative asymptotic efficiency (ARE) and empirically via a simulation study. It is shown that the performance of a test depends on the clusters configurations and on the intracluster correlations. Explicit formulas for optimal weights and a discussion of the impact of omitting a level are provided for 2- and 3-level data.

383 Health Risk and Exposure Assessment ●

Section on Risk Analysis

Wednesday, August 1, 8:30 am–10:20 am

Bayesian Exposure Assessment for Pesticide Intake from Multiple Food Products

❖ Ayona Chatterjee, University of West Georgia, 3201 Post Woods Dr Apt A, Atlanta, GA 30339, *achatter@westga.edu*; Graham Horgan, Biomathematics & Statistics of Scotland; Chris Theobald, University of Edinburgh

Key Words: exposure assessment, latent variable models, bayesian approach, correlated intakes

Pesticide risk assessment involves combining information from consumption and concentration data sets to obtain a distribution for the pesticide intake in a human population. In this work we present a probabilistic, Bayesian approach to modeling the intake of the pesticide Iprodione though multiple food products. We develop latent-variable models that allow for both skewness and large numbers of zeros in the consumption and concentration data. Our consumption model also accounts for correlated intakes. We combine predicted intakes and concentrations from these models to obtain a distribution for individual Iprodione exposure. The use of our probabilistic approach is intended to yield more robust estimates of high percentiles of the exposure distribution than an empirical approach. Bayesian inference is used to facilitate the treatment of data with a complex structure.

The Application of the Benchmark Approach to the Study of the Joint Effect of Mixtures of Air Pollutants

❖ David Svendsgaard, U.S. Environmental Protection Agency, B24301, 109 Alexander Dr, Research Triangle Park, NC 27711, *svendsgaard.david@epa.gov*

Key Words: benchmark risk assessment, chemical mixtures, air pollutants

The aim is to estimate the joint effect of the mixture of five air pollutants (PM10 , O 3 , SO2 , NO2 and CO) for time series mortality/air pollution data sets of large cities in the U.S. In a single city, some of these pollutants can be highly correlated, making it difficult to separate the effects of individual chemicals. We estimate the expected the daily death rates due to meteorology and time components each city using the General Additive Model. Using the Benchmark approach in (Budtz-Jorgensen et al. 2001 Biometrics 57, 698-706), a measure of adjusted death rates is calculated, and the unusually high daily rates are identified. This measure is the dependent variable used for various nonlinear chemical mixture models (Greco et al.1988 Proceedings of the Biopharm. Sect. of ASA). The data are pooled to improve the separation of each pollutant's effects by reducing the colinearity.

Bayesian Analysis of Physiologically Based Pharmacokinetics Modeling of Perchloroethylene in Humans

❖ Junshan Qiu, University of Georgia, 225 Venita drive, Athens, 30605, *qiu@uga.edu*

Key Words: PBPK model, Perchloroethylene, Markov Chain Monte Carlo, Bayesian

Perchloroethylene (Perc) is a widely distributed pollutant. The cancer risks of Perc at low exposures are uncertain. Perc occurs widely in the dry cleaning establishments and also can be found in indoor air. However, the concentrations of Perc are mostly below 1ppm. Therefore, it is very important to assess cancer risks at these low concentrations. A human physiologically based pharmacokinetics (PBPK) model was used to quantify tissue doses of Perc and its key metabolite, Trichloroacetic Acid (TCA) after inhalation exposures. This PBPK model was integrated with a statistical hierarchical model to acknowledge variations due to intra/interindividual variation, measurement error and difference between study methods. One important contribution of this study is to estimate population distributions of PBPK model parameters. Another is to make a dose reconstruction with clinical data.

Parametric Modeling of Melanoma Prognoses: An Alternative to the Cox Regression Model

❖Shouluan Ding, The University of Alabama at Birmingham, NP 2540, 1802 6th Avenue South, Birmingham, AL 35242-3300, *sding@uab.edu*; Huiyi Lin, The University of Alabama at Birmingham; Renee Desmond, The University of Alabama at Birmingham; Seng-jaw Soong, The University of Alabama at Birmingham

Key Words: AFT Models, Melanoma Survival Data, Extended Generalized Gamma Distribution, Parametric models

We investigated the accelerated failure time (AFT) parametric survival model as applied to multivariate prognostic modeling of localized melanoma data with six known potential prognostic factors. Using probability plotting, Cox-Snell residual plots, and likelihood-ratio statistic test methods, we demonstrated that an extended generalized gamma (EGG) survival function was the best fitted parametric model applicable to melanoma prognosis. Using this criteria and the relationship of scale and shape parameters of EGG, all showed that the EGG model adequately served as an alternative model to Cox regression model in accurately predicting survival rates in melanoma. The EGG model provided more efficient parameter estimates compared to the Cox model. It also estimated hazard and survival functions for individualized patient prognosis, making it more useful for clinical interpretation.

A Bayesian Network-Driven Approach for Quantitative Risk Assessment of Foreign Body Injuries in Children

❖Cecilia Scarinzi, Universit‡ degli Studi di Torino, piazza arbarello 8, Torino, 10121 Italy, *scarinzi@econ.unito.it*; Paola Berchialla, University of Torino; Silvia Snidero, S&A S.r.l.; Alexander Stancu, University of Torino; Roberto Corradetti, University of Torino; Dario Gregori, University of Torino

Key Words: Bayesian network, children, foreign body injuries, quantitative risk assessment

Injuries due to foreign body (FB) aspiration/ingestion among children represent a common public heath issue. The aim of this study is to provide a Bayesian Network (BN) that allows for the identification and quantification of risk factors for FB injuries. Combining in a coherent way qualitative and quantitative data, a BN learning algorithm was used to generate a network of the complex relationships among causal factors, both in its topology and its distribution, followed by model verification. Thus, the BN was used for drawing inference on propositions of interest like the probability of injured children to experience complications or hospitalization. Using BN is a new method to evaluate unknown outcomes. This tool of prediction analysis may assist in determining potential hazard of object characteristics. Previous statistical analyses do not allow individual risk analysis.

Confidence Intervals for Quantiles Using Generalized Lambda Distributions

❖Steve Yu Shuo Su, The George Institute for International Health, University of Sydney, Australia, Level 24, 207 Kent Street, Sydney, International 2000 Australia, *ssu@george.org.au*

Key Words: Generalised lambda distributions, quantile confidence intervals, density estimation, risk analysis

Generalized lambda distributions (GLD) can be used to fit a wide range of continuous data. As such, they can be very useful in estimating confidence intervals for quantiles of continuous data. This article proposes two simple methods (Normal-GLD approximation and the analytical-maximum likelihood GLD approach) to find confidence intervals for quantiles. These methods are used on a range of unimodal and bimodal data and on simulated data from ten well known statistical distributions (Normal, Student's T, Exponential, Gamma, Log Normal, Weibull, Uniform, Beta, F and Chi-square) with sample sizes n=10,25,50,100 for five different quantiles q=5%,25%,50%,75%,95%. In general, the analytical-maximum likelihood GLD approach works better with shorter confidence intervals and has closer coverage probability to the nominal level as long as the GLD models the data with sufficient accuracy.

384 Graphs and Networks ● ✪

Section on Statistical Computing, Section on Government Statistics, Section on Statistical Graphics, Section on Statisticians in Defense and National Security

Wednesday, August 1, 8:30 am–10:20 am

Structural Learning of Chain Graphs via Decomposition

❖Zongming Ma, Stanford University, Sequoia Hall 390 Serra Mall, Stanford University, Stanford, CA 94305, *zongming@stanford.edu*; Xianchao Xie, Harvard University; Zhi Geng, Peking University

Key Words: Bayesian network, chain graph, directed acyclic graph, junction tree, structural learning, undirected graph

Chain graphs form a broad class of graphical models for description of conditional independence structures, including both undirected graphs and directed acyclic graphs as special cases. In this paper, we propose that the structural learning of a chain graph can be decomposed into local structural learning related to its decomposed subgraphs. Algorithms for both skeleton recovery and complex arrow orientation are presented on the basis of the decomposition. The decomposition requires conditional independencies but doesn't require the separators to be complete subgraphs. By decomposition, we localize the search for c-separators into small subgraphs, which improves both the efficiency of structural learning and the power of conditional independence tests.

A Statistical Social Network Approach to Computer Network Optimization

❖Hadi Rezazad, 7271 Highland Estates Place, Falls Church, VA 22043, *hrezazad@orchidtech.net*; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University

Key Words: Social Networks, Computer Networks, Centrality, Efficiency, Robustness

Computer network design involves consideration of both fault tolerance and efficiency. In many settings these design issues are antithetic. We review work on using social networks for optimizing LAN networks. We consider vulnerabilities of highly centralized and highly decentralized servers and use evolutionary computing techniques to scale the social network optimization methods to enterprise level networks.

Scale-Free Origin-Destination Flow Statistics Estimation in a High-Speed Network

❖Aiyou Chen, Alcatel-Lucent, 2c358, 600 Mountain Ave, Murray Hill, NJ 07974, *aychen@alcatel-lucent.com*; Jin Cao, Bell Labs; Tian Bu, Bell Labs

Key Words: network traffic, data stream, flow statistics

Many aspects of network management, such as network provisioning, and traffic anomaly detection, rely on knowing the traffic flow statistics between pairwise network nodes. However, rapid growth of networks in size and bandwidth makes obtaining such statistics very challenging. In this paper, we propose efficient streaming algorithms to generate compact on-line statistical digests of traffic, from which likelihood based estimates are derived. The new approaches are scale-free, i.e., the accuracy essentially does not depend on the unknown traffic parameters. We perform a thorough theoretical and simulation analysis of our algorithms and evaluate them using a traffic matrix from a top-tier Internet Service Provider. These analyses demonstrate that the new algorithms yield more accurate estimates than existing approaches, and are robust to heterogeneous and dynamic network conditions.

Style of Author-Coauthorship Social Networks: Statisticians of Prominent U.S. Universities

❖ Walid Sharabati, George Mason University, 403 Hillsboro Dr, Silver Spring, MD 20902, *wsharaba@gmu.edu*; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University

Key Words: social networks, coauthorship, cliques, blockmodeling, MDS clustering, preferential attachment

In the past 20 years social networks have been used to analyze relations and ties among individuals of the same network and similarities between different networks in an attempt to obtain a better understanding on how societies interact. One of the applications of social networks is the author-coauthor networks also known as the citation networks. This branch of social networks tries to answer the question of "who-wrote-with-who" and with what frequency. It also investigates other important features such as cliques, structural equivalence, MDS and CONCOR clustering. Social networks can be treated as directed graphs in which actors (individuals) are represented by vertices (nodes) while interactions between actors are represented by edges (ties) which may have weights. In this paper, we study in depth coauthorship social networks of statisticians from prominent U.S. universities.

Partitioning the Sample Space on Five Taxa for the Neighbor Joining Algorithm

❖ Ruriko Yoshida, University of Kentucky, 805A Patterson Office Tower, Lexington, KY 40506, *ruriko@ms.uky.edu*; Kord Eickmeyer, Institute fir Informatik, Humboldt-Universitat zu Berlin

Key Words: neighbor joining, phylogenetic diversity, phylogenetic trees

The Neighbor Joining Algorithm is a recursive procedure for reconstructing trees that is based on a transformation of pairwise distances between leaves and it is widely used because of its accuracy and computational speed. In this talk, we will analyze the behavior of the Neighbor Joining algorithm on five taxa and we will show that the partition of the sample (data) space for estimation of a tree topology with five taxa into subspaces, within each of which the Neighbor Joining algorithm returns the same tree topology. A key of our method to partition the sample space is the action of the symmetric group S_5 on the set of distance matrices by changing the labels of leaves. The method described in this paper can be generalized to trees with more than five taxa.

Destratified Importance Sampling

❖ Joseph Blitzstein, Harvard University, Statistics Dept, 1 Oxford St, Cambridge, MA 02138, *blitzstein@stat.harvard.edu*

Key Words: importance sampling, networks, graphs, sequential Monte Carlo

Importance sampling is often employed for estimation in large state spaces. In some situations though, the trial distribution is unknown or computationally difficult. We describe an approach to such problems, in which the state space is enriched to a larger space, and then importance sampling on the larger space can be used to induce efficient importance sampling estimates for the original space. As an application, we use the method to estimate average geodesic distances in various network models.

Process Modeling To Explain Ordinal Categorical Data

❖ Simone Gray, Duke University, Institute of Statistics Decision Scien, Box 90251 214 Old Chemistry Building, Durham, NC 27708, *simone@stat.duke.edu*; Alan E. Gelfand, Duke University

Key Words: Ordinal, Categorical, Data

Let us suppose that all classifying variables in a multi-way contingency table are ordinal. Then expected cell counts in the table can be viewed as driven by an intensity surface over R^m. We propose to model this surface as a realization of an m-dimensional spatial process. We can envision two versions - an ordinal classification associated with known scale and known categorical intervals, and an ordinal classification associated with a conceptual latent scale and unknown intervals. Moreover, we may have some of each, as well as nominal categories. Benefits of this approach include flexible modeling for the joint probabilities in the table and interpolation to other ordinal classifications for the variables. We detail modeling for such settings and the associated computation to fit these models. We illustrate with both simulated data and real data from the North Carolina birth records.

385 Creating Cognitive Thinking Opportunities in Introductory Statistics Classes

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Wednesday, August 1, 8:30 am–10:20 am

Shades of Gray: Ambiguity Tolerance and Inferential Thinking

❖ Robert Carver, Stonehill College, 39 Woodland Street, Sharon, MA 02067, *robert.carver@comcast.net*

Key Words: Statistical thinking, Education, Introductory statistics

Students sometimes express frustration with the uncertainty that remains after solving a problem of statistical inference: for instance, even with a "correct" solution we don't know if a confidence interval includes the parameter of interest. This paper reports on a continuing empirical investigation of the relationship between Ambiguity Tolerance and students' facility in developing the skills of inferential reasoning. On the one hand, inferential methods provide an approach for reasoning about ambiguous situations and therefore might have special appeal for students frustrated or agitated by ambiguity. On the other hand, ambiguity intolerance might present a roadblock to serious engagement in the study of inferential methods. This research focuses explicitly on ambiguity tolerance as an explanatory or moderating factor in learning to apply the techniques of inference.

Using Student-Generated Survey Data To Reinforce Basic Concepts of Introductory Statistics

❖ Elizabeth Walters, Loyola College in Maryland, 4501 N Charles St, Baltimore, MD 21210, *ewalters@loyola.edu*

In an introductory statistics course, hands-on student learning is essential for maximum success. Experience in recent years has shown that introducing real-world data into the classroom is also an effective teaching tool. This paper will discuss the implementation and results of an approach to combine these two tools to teach the basic concepts of introductory statistics through use of real data collected by students. Students will collect data by use of random sampling techniques and sound survey methods.

These data will then be used in assignments throughout the semester to reinforce concepts such as correlation, regression, and chi-square tests.

How Much Mathematics Is Appropriate in a Modern Statistics Course?

❖Brian E. Smith, McGill University, 1001 Sherbrooke St West, Montreal, QC H3A 1G5 Canada, *brian.smith@mcgill.ca*

Key Words: mathematics, formula, software, concept

This presentation will examine the presenter's experiences in shifting from a traditional formula driven business statistics course to a format with greater emphasis on a conceptual, largely computer based, approach. The focus of the new course is on presenting major statistical concepts via a combination of visualization and interpretation of computer output. An important question arises with regard to the reduction of the quantitative component of the course, namely, how little is too little? In the author's experience there is a subset of students who feel disadvantaged when formulas are glossed over and when numerical results are produced via the black box effect of having them produced by a software package. This dilemma will be examined from the perspective of a statistics instructor with a strong mathematical background and a commitment to teaching an excellent statistics course.

Misconceptions in Interpreting Center and Variability in Graphical Representations

❖Linda Cooper, Towson University, Mathematics Department, 8000 York Road, Towson, MD 21252-0001, *lcooper@towson.edu*; Felice Shore, Towson University

Key Words: misconceptions, graphs, introductory statistics, mean, median, variability

This paper will identify and discuss misconceptions that students have in making judgments of center and variability when data is presented graphically. An assessment addressing interpreting center and variability in histograms and stem-and-leaf plots was administered to undergraduates enrolled in introductory statistics courses, including prospective elementary and secondary teachers. In particular, discussions will focus upon comparing the variability of two data sets of common range represented by bell-shaped histograms on common scales, and in comparing the relative location of the mean and median on a histogram from skewed data. When the exact values of the data were available in stem-and-leaf plot representation, students were more successful at computing the mean and median values. Results of the assessment questions and subsequent implications for instruction will be discussed.

College Students' Intuitive Reasoning About Variation of Sample Mean

❖Carl Lee, Central Michigan University, PE 109 Department of Mathematics, Mt Pleasant, MI 48859, *carl.lee@cmich.edu*; Maria Meletiou-Mavrotheris, Cyprus College

Key Words: Sampling distribution, Statistical reasoning, Variation

Variation is one of the most important concepts in Introductory Statistics. This article studies students' intuitive reasoning about variation of sample mean prior to taking a basic statistics course. Over 200 students are involved in the study. A sequence of five open-ended questions based a real problem scenario are given. Students are asked to answer each question and the reasons. Their responses and reasons are analyzed. The results suggest that their intuitive reasoning about variation of sample mean is better than we expected. Some possible explanations are (a) the design of the question sequence may help them to connect the variations of individual observations and sample mean, (b) the problem scenario may help students in building their reasoning, and (c) without a formal introduction of the subject gives students a flexible and non-threatening critical thinking space.

Observational Creativity: Statistical Reasoning and the Dead Sea Scrolls

❖Donald Bentley, Pomona College, 1826 Roanoke Rd, Claremont, CA 91711, *dbentley@pomona.edu*

Key Words: Observational Creativity, Statistics in Humanities, Dead Sea Scrolls

Traditionally, statistical reasoning has been considered as a tool to be applied to problems arising in disciplines within the physical, biological, and social sciences. Consequently, statistics courses are taught using examples of applications to and data from fields within these areas. Little time or effort is spent encouraging either students or teachers to look outside these areas for examples of problems where statistical reasoning might contribute to a solution. But disciplines within non-traditional areas, in particular the humanities, can be fertile ground for the application of statistical reasoning. In this paper we will present in a non-technical way three statistical problems that grew out of data contained in the Dead Sea Scroll collection, questions which were motivated by research in Biblical Scholarship for which statistical reasoning might help find an answer.

Discovering the Underlying Themes of Statistical Models

❖Steve Wang, Swarthmore College, 500 College Ave, Dept of Math and Stat, Swarthmore, PA 19081, *scwang@swarthmore.edu*

Key Words: education, teaching, model, uniform distribution, paleontology, learning activity

Statisticians are fond of models, but communicating this abstract concept to introductory students can be challenging. Here I describe a class lecture and activity designed to help students understand how a statistical model represents (and necessarily simplifies) reality. In class, I present three examples that at first seem unrelated. The first is the familiar "Nazi tank" activity, in which students try to devise estimators for the number of tanks possessed by the Germans in World War II. Next, I present an activity in which students estimate how many people are in the classroom. Finally, I present an example from my own research on estimating the extinction time of fossil species. I challenge students to find the common theme in these apparently unrelated examples, thereby leading them to discover the idea of a statistical model through active learning.

386 Space-Time Modeling

Section on Statistics and the Environment, General Methodology, Section on Statisticians in Defense and National Security

Wednesday, August 1, 8:30 am–10:20 am

Characterizing the Dependence Structure of Space-Time Processes Using Computer-Model Output and Sparse Observations

❖Candace Berrett, The Ohio State University, 3504 Prestwick Ct, Columbus, OH 43220, *cberrett@stat.ohio-state.edu*; Catherine A. Calder, The Ohio State University; Tao Shi, The Ohio State University

Key Words: Atmospheric science, Environmental science, Remote sensing, Spatial statistics, Time Series

Characterizing the dependence structure of a space-time process from remote-sensing data can be a difficult task. Large amounts of missing observations, as well as systematic variation in the quality of observations, do not allow important features such as nonstationarity and anisotropy to

be readily identified. In this talk, we discuss how output from numerical models can be used to guide space-time modeling of a process for which only limited data are available. In particular, we focus on modeling the distribution of carbonaceous aerosols over mainland Southeast Asia using data collected by the MISR and MODIS instruments onboard the Terra and Aqua satellites and output from the aerosol transportation simulator MOZART.

Bayesian Hierarchical Models for Extreme Values Observed over Space and Time

❖ Huiyan Sang, Duke University, Rm 214 Old Chem Bldg, Durham, NC 27705, *hs37@duke.edu*; Alan E. Gelfand, Duke University

Key Words: generalized extreme valued distribution, Spatial random effects, MCMC Langevin algorithm, Markov random field, Temperature surface

We propose a hierarchical Bayesian approach for modeling a collection of spatially-referenced time series of extreme values. We assume that the observations follow Generalized Extreme Value (GEV) distributions where the dependence is captured using the transformed spatial Gaussian process and the spatially dependent locations and scales. The models can be fitted using a Markov Chain Monte Carlo (MCMC) Langevin algorithm to enable inference for parameters and to provide spatio-temporal predictions. We fit the models to a set of weather station temperature data collected over a 50-year period for the Cape Floristic Region in South Africa.

Semiparametric Zero-Inflated Poisson Models for Spatio-Temporal Environmental Processes

❖ Ali Arab, University of Missouri-Columbia, 55 Broadway Village Dr, Apt G, Columbia, MO 65201, *aa5vf@mizzou.edu*; Christopher K. Wikle, University of Missouri-Columbia; Scott Holan, University of Missouri-Columbia

Key Words: spatio-temporal models, zero-inflated models, environmental processes, Bayesian, multivariate

Environmental studies often include multivariate count processes characterized by interactions at different spatial and temporal scales, as well as, excess zeros resulting from the underlying natural phenomena and/or imperfect detection and sampling procedures. However, lack of multivariate discrete distributions that allow for complex dependence structures and excess zeros, and computational difficulties involved in application of such distributions in a joint modeling framework, necessitate the development of efficiently parameterized hierarchical Bayesian models. Furthermore, a semiparametric modeling approach provides a more realistic framework in which to account for possible non-linear relationships among variables in real-world situations. We develop hierarchical Bayesian semiparametric zero-inflated models for multivariate spatio-temporal environmental count processes.

A Spatial-Temporal Point Process Model for Nowcasting Radar Reflectivities

❖ Yong Song, University of Missouri-Columbia, 701 S Providence Rd, Department of Statistics, Apt 4D, Columbia, MO 65203, *takkisong@gmail.com*; Christopher K. Wikle, University of Missouri-Columbia

Key Words: Spatial-temporal, Point process, Movement, Nonlinear, Bayesian, Nowcasting

Short-term forecasts of heavy rainfall involve forecasting the movement of the radar-based estimates of precipitation over fairly short time-scales. The tradition methods of nowcasting weather radar reflectivities are based on linear extrapolation approaches and physical (deterministic) approaches. These methods have limitations. The deterministic approach is difficult to apply in near real time, and does not account for uncertainties in dynamical parameterizations. Extrapolation methods often fail due to the nonlinear nature of the process evolution. In this project, the authors employ a spatial-temporal point process model for short-term forecasts of heavy rainfall events. A nonlinear structure and the dependence between the rainfall cells have been considered. Bayesian approaches have been applied in developing this model.

Identification of Pollution Source Directions via MCMC

❖ Basil Williams, Brigham Young University, 414 River Meadow Drive, Alpine, UT 84004, *basil.williams@gmail.com*; William Christensen, Brigham Young University

Key Words: pollution source apportionment, PSA, circular data, bayesian regression, directional data

Pollution source apportionment (PSA) is the practice of identifying and describing pollution sources and their contributions. PSA frequently requires the identification of source directions, often as a post-analysis check to ensure that the contribution estimates are reasonable. This paper develops a method of identifying source directions using Bayesian regression. MCMC is used to evaluate the complex relationship among observed pollutant concentrations, available meteorological information, and unknown source direction parameters. The method is flexible enough to identify multiple source directions for cases in which a species or source type of interest is emitted at more than one location.

Bias in the Estimation of Self-Exciting Point Process Models for Earthquake Occurrences

❖ Annie Chu, University of California, Los Angeles, 820 S Ynez Avenue, Monterey Park, CA 91754, *chea@stat.ucla.edu*; Frederic Paik Schoenberg, University of California, Los Angeles; Alejandro Veen, IBM T.J. Watson Research Center

Key Words: earthquakes, epidemic-type aftershock sequence models, ETAS models, space-time point process models, branching process models, bias in MLE

The space-time Epidemic-Type Aftershock Sequence (ETAS) model developed by Ogata (1998) is a branching point process model having been successfully used to characterize certain earthquake catalogs. However, many modern earthquake catalogs are missing many events, especially those of small magnitude. Hence most analyses institute a lower magnitude threshold before fitting a model such as ETAS. The impact of the magnitude cutoff on the branching ratio has been investigated by Sornette and Werner (2005). We investigate the impact of the magnitude threshold on the bias in maximum likelihood estimates of the parameters in the ETAS model. Our investigations make repeated use of an EM method to obtain stable MLE for branching point process models, developed by Veen (2006). Results are demonstrated with simulations and use a catalog compiled by the Southern California Earthquake Center (SCEC).

An Application of Voronoi Diagrams to Modeling Earthquake Catalogs

❖ Christopher Barr, University of California, Los Angeles, 8117 Math Science Building UCLA, Box 951554, Los Angeles, CA 90095-1554, *chrisbarr@stat.ucla.edu*; Frederic Paik Schoenberg, University of California, Los Angeles

Key Words: Voronoi Diagrams, Tapered Pareto Distribution, Earthquakes, Spatial Statistics, Stochastic Geometry

Following a brief summary of the use of Voronoi diagrams for problems in such diverse fields as astronomy, biology and operations research, some new applications in describing the spatial structure of earthquake epicenters will be introduced. Seismic moments and other properties of earthquakes have long been modeled by the Pareto distribution. Here, a twenty-year catalog of Southern California earthquakes is analyzed, and evidence is provided which suggests numerous properties of Voronoi cells arising from tessellated epicenters follow the tapered Pareto distribution. The findings relate to previous results on the tapered Pareto distribution of earthquake sizes and aftershock distances by a variety of authors including Kagan, Vere-Jones, and Ogata. Additionally, recent work using the tapered Pareto to model wildfire burn regions potentially suggests even broader applications.

387 Bayesian Methods and Applications in Epidemiology ●

Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences, Section on Health Policy Statistics

Wednesday, August 1, 8:30 am–10:20 am

HIV Dynamics and Natural History Studies: Joint Modeling with Interval-Censored Event Times and Sparse Longitudinal Data

❖ Li Su, Brown University, 121 South Main Street, 7th FL, Providence, RI 02912, *lisu@stat.brown.edu*; Joseph W. Hogan, Brown University

Key Words: HIV/AIDS, HCV, Antiviral treatment, Viral dynamics, Penalized splines, Bayesian methods

The effect of HCV coinfection on post-HAART HIV dynamics has drawn considerable attention of researchers recently. In this article we develop a model for characterizing post-HAART HIV dynamics and their association with HCV coinfection using sparse cohort data from natural history studies. We propose a joint model for three processes: time from enrollment to HAART initiation, time from HAART initiation to viral suppression, and evolution of CD4 counts relative to viral suppression time. Inference is accomplished using a fully Bayesian approach. Interval-censored event times are modeled semiparametrically by Dirichlet process priors. Bayesian penalized splines are used for modeling population-level and individual-level CD4 count profiles. We use proposed methods and data from HIV Epidemiology Research Study to investigate the effect of HCV coinfection on the response to HAART.

Covariate-Adjusted Constrained Bayes Predictions of Random Intercepts and Slopes

❖ Robert Lyles, Emory University, Dept of Biostatistics, 1518 Clifton Rd NE, Atlanta, GA 30322, *rlyles@sph.emory.edu*; Renee Moore, University of Pennsylvania; Amita K. Manatunga, Emory University; Kirk Easley, Emory University

Key Words: Exposure modeling, HIV, Prediction, Random effects

Standard prediction of random effects under the mixed linear model takes an empirical Bayesian approach to produce estimates of their posterior mean given the data. While such shrinkage estimates optimize mean square error, the narrow spread of the resulting predictor distribution can be undesirable, e.g., when the objective is to classify subjects relative to a threshold. The constrained Bayes method provides a reduced shrinkage alternative. We examine this approach for predicting random effects, with particular attention to the handling of covariates. We find the general method of Ghosh (1992) to be flexible and comparable with a direct implementation of constraints, suggesting potential for its incorporation into common software for mixed linear models. We provide an example aiming to predict CD4 cell counts among HIV-infected children at the time of Class A disease diagnosis.

Family-Based Case-Control Studies

❖ Li Zhang, The Cleveland Clinic, 6809 Mayfiled Road, Apt 555, Cleveland, OH 44124, *zhangl3@ccf.org*; Bhramar Mukherjee, University of Michigan

Key Words: Dirichlet Process, hierarchical pedigree, conditional likelihood, mixed effects model, random effects model

As a compromise between linkage studies and population-based case-control studies, family-based association designs have received great attention recently due to their potentially higher power to identify complex disease genes and their robustness in the presence of population substructure. Based on a two-level mixed effects model which allows to estimate environmental effects while accounting for varying genetic correlations among family members and adjusting for ascertainment by conditioning on the number of cases in the family, we propose a full Bayesian alternative to build in a hierarchical pedigree structure and assuming priors on the random effects which offers a more appealing alternative. We provide a general framework for Bayesian analysis for the random effects model where a nonparametric Dirichlet Process prior is specified for the random effects.

Association Test Using Pedigree and Multilocus Data with Missing Genotypes and Population Structure

❖ Jie Ding, The Ohio State University, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *ding@stat.ohio-state.edu*; Shili Lin, The Ohio State University

Key Words: Association, Pedigree data, Missing genotypes, Population structure

Association test using pedigree data is one way of mapping disease susceptibility genes. Previously we have proposed the Monte Carlo pedigree disequilibrium test (MCPDT) to test for association using pedigree data with missing genotypes by averaging over a Monte Carlo sample of those genotypes generated by conditioning on known genotypes. Here we extend MCPDT by taking into account the population structure in a Bayesian framework. An MCMC algorithm is designed to infer population structure from many markers and simulate missing genotypes at the testing marker loci. This algorithm can be applied to both single marker locus and haplotypes consist of multiple loci.

Regression Models for Paired Censored Exponential and Gamma Distributions

❖ Michael P. Jones, Macquarie University, Psychology Department, C3A, North Ryde, 2080 Australia, *mike.jones@psy.mq.edu.au*; Don R. McNeil, Macquarie University

Key Words: Bivariate censoring, Survival, Models

Background: Models for nonindependent observations are well-studied. While some situations are well covered, such as repeated failures, others, such as twin studies, are not. Aim: Develop models suitable for paired observations where one or both pair members (y, z) might be right-censored. Method: For both Exponential and Gamma distributions we consider 1) a Bayesian approach eliminating nuisance parameters by assuming a non-informative prior and 2) derivation of distributions of ratios ($r=y/z$). Likelihood functions are derived and maximized numerically. Results: (1) requires only a single parameter to represent the within-pair dependence while the ratio approach eliminates nuisance parameters altogether. Simulation studies show the asymptotic Normality of the MLEs for both (1) and (2) but suggest greater statistical power from (1). Conclusion: The proposed methods are simple.

Modeling a Cross-Sectional Response Variable with Longitudinal Predictors: An Example of Pulse Pressure and Pulse Wave Velocity

❖Veena Shetty, National Institutes of Health, NIA, Labortaory of Cardiovascular Sciences, 5600 Nathan Shock Drive, Baltimore, MD 21224, *shettyv@mail.nih.gov*; Christopher Morrell, Loyola College in Maryland; Samer S. Najjar, National Institutes of Health

Key Words: Mixed effects, Linear regression, Area under the curve

We wish to model pulse wave velocity (PWV) as a function of longitudinal measurements of pulse pressure (PP) at the same and prior visits at which the PWV is measured. A number of approaches are compared. First, we use the PP at the same visit as the PWV in a linear regression model. Next, the PP at the visit as well as the rate of change in PP are used to model PWV. An approach for using the longitudinal PP data is to obtain a measure of cumulative burden, the area under the PP curve (AUC). This AUC is used as an explanatory variable to model PWV. Next, a two stage process is applied. The longitudinal PP is modeled using a linear mixed-effects model. Then the PP and rate of change in PP at the visit where the PWV is measured are estimated from the model and used in the regression model to describe PWV. Finally, a joint Bayesian model is constructed similar to the two-stage model.

ROC Analysis for Longitudinal Disease Diagnostic Data Without a Gold Standard Test

❖Chong Wang, Cornell University, 210 Malott Hall, Department of Mathematical Sciences, Ithaca, NY 14850, *cw245@cornell.edu*; Bruce W. Turnbull, Cornell University; Yrjö T. Gröhn, Cornell University; Søren S. Nielsen, The Royal Veterinary and Agricultural University

Key Words: Bayesian, Change-point Models, Markov chain Monte Carlo methods

We develop a Bayesian methodology based on a latent change-point model to estimate the ROC curve of a diagnostic test for longitudinal data. We consider the situation where there is no perfect reference test, i.e., no "gold standard." A change-point process with a Weibull-like survival hazard function is used to model the progression of the hidden disease status. Our model adjusts for the effects of covariate variables, which may be correlated with the disease process or with the diagnostic testing procedure, or both. Markov chain Monte Carlo methods are used to compute the posterior estimates of the model parameters that provide the basis for inference concerning the accuracy of the diagnostic procedure. We discuss an application to an analysis of ELISA scores in the diagnostic testing of paratuberculosis (Johne's disease) for a longitudinal study with 1997 dairy cows.

388 Topics in Survey and Census Redesigns ●

Section on Survey Research Methods

Wednesday, August 1, 8:30 am–10:20 am

Strategic Directions for the 2011 Canadian Census of Population

❖David Dolson, Statistics Canada, 15 O RH Coats Building, 100 Tunneys Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *ddolson@statcan.ca*; Normand Laniel, Statistics Canada

Key Words: Cost Reduction, Automation, Data Quality, Data Collection

For the 2011 Census, Statistics Canada will be implementing a number of refinements and further changes, building on the major updating of methodology undertaken in 2006. This paper will discuss major drivers, strategic issues and planned changes. Foremost is cost reduction and reduced reliance on a large staff for field operations while maintaining data quality. Significant changes planned for updating of the address register, data collection (e.g., targeting an Internet take-up rate of 40% as well as an extension of mail-out), non-response follow-up and follow-up for edit failures will be reviewed. Plans for a variety of other changes (e.g., use of paradata in non-response adjustment, larger blocks of questions in imputation, Ridge-type regression for weighting, improved variance estimator) will also be presented.

Alternative Allocation Designs for a Highly Stratified Establishment Survey

❖Ernest Lawley, Bureau of Labor Statistics, 2 Massachusetts Ave NE, #4985, Washington, DC 20212, *lawley.ernest@bls.gov*

Key Words: Neyman allocation, probability proportional to estimated size (PPES), power allocation, establishment surveys, stratified surveys

The primary objective of Occupational Employment Statistics Survey, conducted by U.S. Bureau of Labor Statistics in partnership with the 50 States and District of Columbia, is to measure occupational employment and wages at the very detailed level of Metropolitan Statistical Areas (MSA) crossed by over 300 industries. That is, how many people are employed in one of the 800 Standard Occupational Codes (SOC) and what are the mean occupational wages for each industry by MSA. A given sampling frame contains about 175,000 non-empty MSA-by-industry cells. The occupational employment and wage estimates are also required at various aggregated levels of geography and industry. This study examines alternative sample allocation designs for a highly stratified population that deals with multiple issues such as establishment employment size and occupational diversity and variability.

Sampling with Uncertain Frame Counts: Challenges in Sampling Head Start Children for the FACES Study

❖Barbara L. Carlson, Mathematica Policy Research, Inc., PO Box 2393, Princeton, NJ 08543-2393, *bcarlson@mathematica-mpr.com*

Key Words: FACES, Head Start, oversampling, frame problems

The 2006 Head Start Family and Child Experiences Survey (FACES) involved four stages of sampling: Head Start programs, centers, classrooms, and children. Eligible children were those who were one or two years away from kindergarten and were new to Head Start in the fall of 2006. Because only a list of Head Start programs was available as a sampling frame, we relied on selected programs to provide lists of centers, and relied on selected centers to provide lists of classrooms and eligible children. To accommodate when the lists were provided, sample selection at each level was conducted on a rolling basis. We will describe the challenges in implementing a sampling strategy that met the sample design goals, including an oversample of children who were two years away from kindergarten, and one that was flexible enough to adapt to the actual child counts when they were lower than estimated.

Sample Design Research for the National Home and Hospice Care Survey

❖Karen Davis, National Center for Health Statistics, 3311 Toledo Road, Rm 3213, Hyattsville, MD 20782, *kedavis@cdc.gov*; Iris Shimizu, National Center for Health Statistics

Key Words: sample survey design, design effects

The National Home and Hospice Care Survey (NHHCS) is a nationally representative sample survey of home health and hospice care agencies, their patients, and staff conducted by the National Center for Health Statistics (NCHS). Estimates from the NHHCS provide important information on the users and providers of long term care, which is valuable for policymakers in the U.S. as large segments of the population continue to age. The NHHCS has been conducted six times since 1992 (1992, 1993, 1994, 1996, 1998 and 2000). The NHHCS was not fielded in 2002 or 2004 to allow time to conduct survey and sample design developmental work that would facilitate future survey redesign efforts. This paper describes the research objectives, 2000 sample design, design changes and ongoing research activities. A redesigned NHHCS is scheduled to be fielded in summer 2007.

Sample Design of the Canadian Health Measures Survey

❖Sarah M. Dion, Statistics Canada, 41 Hilda Street Unit 303, Ottawa, ON K1Y4T4 Canada, *sarahmaude.dion@statcan.ca*

Key Words: Health Survey, Multistage Sample, Direct Measures, Vector of Selection, Sub sampling

Statistics Canada, in partnership with Health Canada and the Public Health Agency of Canada, is initiating the Canadian Health Measures Survey (CHMS) beginning in the spring of 2007. The CHMS aims to overcome the limitations of existing health-monitoring information by directly measuring health indicators from a nationally representative sample of 5,000 Canadians aged 6 to 79. The respondents will be asked to complete an in-home health questionnaire and then travel to a clinic to have physical health measurements taken by professionals. A multistage sample was developed to meet the objectives and logistics of the survey. The presentation will provide an overview of the CHMS with emphasis on the sample design. The effectiveness of the design will also be examined, using data collected for the first two sites.

Sampling and Estimation Issues in the Redesign of the Canadian Survey of Household Spending

❖Christian Nadeau, Statistics Canada, RH Coats Building Floor 16, Tunney s Pasture, Ottawa, ON K1A 0T6 Canada, *christian.nadeau@statcan.ca*

Key Words: Expenditure survey, Diary, Recall interview, Weighting

The Survey of Household Spending is conducted by Statistics Canada at the beginning of each year through recall interviews. A major redesign has been undertaken in order to spread the collection of data over the year, to adapt the reference periods to the recall capacity of the respondents and to integrate the Food Expenditure Survey. A two-week diary will allow the collection of frequent expenditure data while less frequent data will be collected through an interview for recall periods varying from one to twelve months. The implementation of this approach raises a number of sampling and estimation issues. Among these are the combination of the monthly samples in order to produce coherent annual estimates from the interview and the diary data, and the adjustment of the estimates to the calendar year. Different options to be tested on pilot survey data will be discussed in this paper.

Sample Redesign for the FDIC's Asset Valuation Review

❖David W. Chapman, Federal Deposit Insurance Corporation, 7206 Quantum Leap Lane, Bowie, MD 20720, *dachapman@fdic.gov*

Key Words: Optimum allocation, Stratum boundaries, Certainty criteria, Stratified sampling

When an FDIC-insured financial institution is in danger of failing, the FDIC has to assess the value of the institution's loan portfolio in a short period of time, as part of an Asset Valuation Review (AVR). Based on the AVR, price tags are given to various loan pools (types), which are used to sell the bank's loans to other institutions, if the bank fails. Because of the large number of loans in most portfolios, sampling must be used to estimate the value of the loan pools. The basic design is a stratified random sample, with strata defined by size (loan book value). The definition of strata (including a certainty stratum) and the derivation of the total and stratum sample sizes have to be automated because of the limited time available for the AVR. The current design is being revised to improve sampling efficiency. The challenges and features of the revised design will be discussed.

389 Survey Estimation and Inference ●

Section on Survey Research Methods

Wednesday, August 1, 8:30 am–10:20 am

Design Effects of Sampling Frames in Establishment Surveys of Small Populations

Monroe Sirken, National Center for Health Statistics; ❖Iris Shimizu, National Center for Health Statistics, 3311 Toledo Rd., Room 3123, Hyattsville, MD 20782, *IShimizu@cdc.gov*

Key Words: Establishment Sampling Frames, Small Domain Estimation, Network Sampling

This paper compares the design effects of sampling frames in three establishment surveys, E1, E2, E3, of the utilization of establishment services by small populations. Surveys E1 and E2 use population survey-generated sampling frames that list establishments which have transactions with people in household sample surveys. In E1, the population survey-generated sampling frame lists all establishments that have transactions with people in the household sample survey. In E2, the population survey-generated frame lists the subset of establishments that have transactions with the people in the household sample survey that belong to the small population of interest. Survey E3 uses a complete establishment frame that lists all establishments.

Variance Estimation for Statistical Matching Using Multiple Imputation

❖James Reilly, University of Auckland, 5 79 Beresford St West, Freemans Bay, Auckland, 1011 New Zealand, *reilly@stat.auckland.ac.nz*

Key Words: variance estimation, statistical matching, multiple imputation, data fusion, media research

Statistical matching is a method for combining survey databases on distinct topics that is widely used in microsimulation modeling and media research. Several techniques have been used to calculate sampling variances for statistically matched results, including multiple imputation (Rassler 2002), generalized variance functions (Reilly 2000) and resampling methods such as random groups and the jackknife (Ingram et al. 2000; Reilly 2003), but little work has been published comparing these methods. A variant of multiple imputation proposed by Robins and Wang (2000) is newly applied to statistical matching, and the resulting variance estimates are compared with those from traditional multiple imputation and other methods. A New Zealand media research service provides an illustrative application of these techniques.

Using Bootstrap Variance Calculations for a Survey with a Simple Design: The Case of the 2005 National Survey of the Work and Health of Nurses

❖Gildas Kleim, Statistics Canada, 18-A R.H. Coats building, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *gildas.kleim@statcan.ca*; Yves Bélanger, Statistics Canada

Key Words: Variance, Bootstrap, Simple Design

Rao and Wu (1988) proposed a bootstrap method aimed for use in complex sample designs involving stratification, multiple stages and unequal probability sampling. They also presented a bootstrap procedure for stratified simple random sampling without replacement (WOR). Even though formulas for variance estimates for such a simple design are well known, it may happen nevertheless that clients request bootstrap weights to be produced. The usual bootstrap procedure used by Statistics Canada, and referred to as the Rao-Wu bootstrap, assumes sampling with replacement (WR) and thus leads to an overestimation of the variance for WOR sampling, especially when the stratum sampling fractions are high. In this paper, we describe and evaluate an implementation of the bootstrap procedure for a simple design using the 2005 National Survey of the Work and Health of Nurses.

Balanced Repeated Replication Variance Estimates from the Medical Expenditure Panel Survey 1996–2003 Household Component

❖ Dan Liao, University of Maryland, 404 Ridge Rd Apt 12, Greenbelt, MD 20770, *ashley_bnu@hotmail.com*

Key Words: single PSU stratum, balanced repeated replication variance estimation, pooled stratification

In this report, we investigate Balanced Repeated Replication (BRR) variance estimates in the Medical Expenditures Panel Survey (MEPS), Household Component (HC), 1996–2003, for several crucial medical expenditure variables by comparing the results from the data of individual years to the ones from design-based alternatives and the corresponding Taylor linearization variance estimates. For longitudinal analyses of MEPS-HC data, a pooled stratification was created for multiple years to increase the sample size and statistical power. However, in this pooled stratification, some single primary single unit (PSU) strata exist in some individual years which make the design improper for BBR. Three different ways of solution are proposed to adjust this pooled stratification design and the BRR variance estimates of the adjusted designs are tested and evaluated at the end.

Application of the Bootstrap Method in the International Price Program

❖ Patrick Bobbitt, Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Washington, DC 20212, *bobbitt_p@bls.gov*; Steven P. Paben, Bureau of Labor Statistics; Moon Jung Cho, Bureau of Labor Statistics; James A. Himelein Jr., Bureau of Labor Statistics ; Te-Ching Chen, Bureau of Labor Statistics; Lawrence Ernst, Bureau of Labor Statistics

Key Words: Variance estimation, Bias estimation, Certainty Sampling Units, Collapsing

The International Price Program (IPP) collects data on United States trade with foreign nations and publishes monthly indexes on the import and export prices of U.S. merchandise and services. The IPP employs a three stage PPS design in which establishments, then broad product categories traded within establishments, and finally items within a category, are selected. Certainty selections can occur in the first two stages. We present three variations of the bootstrap rescaling method adapted to the IPP sample design: 1) sampling at the first stage, treating certainty units as probability units, 2) sampling that allows for certainties, and 3) a procedure that extends the previous method by collapsing single item strata. Finally, we compare the precision and bias of the three approaches by simulating 1000 samples of a simulated universe using the IPP sampling methodology.

390 Model Specification, Estimation, and Evaluation ●

Social Statistics Section

Wednesday, August 1, 8:30 am–10:20 am

A New Model for Interaction Effects

❖ David Weakliem, University of Connecticut, 344 Mansfield Rd, Storrs, CT 06269-2068, *weakliem@uconn.edu*

Key Words: interaction, general linear model

Interaction effects are often of interest in the social sciences. When the variables involved have a moderate or large number of categories, imposing some structure on the interactions makes results easier to interpret and allows more powerful tests against the additive model. Moreover, different hypotheses about social processes imply different forms of interaction effects. This paper discusses some structured models for two-way interaction effects in the general linear model. The idea of competition among variables (for example, ethnicity and social class) has been important, but is not adequately represented by any existing model for interactions. I propose a model representing the interaction resulting from competition and apply it to a number of examples. The paper also discusses the relationship of the proposed model to several existing models for square tables.

Linear Information Models: A Prototype

❖ Philip E. Cheng, Academia Sinica, 128 Academia Rd, Nankang, Taipei, 115 Taiwan, *pcheng@stat.sinica.edu.tw*; Jiun W. Liou, Academia Sinica; Michelle Liou, Academia Sinica; John Aston, Academia Sinica

Key Words: Information Models, Log-linear Models, Model Selection, Mutual Information

Relative entropy identities yield basic log-likelihood decompositions for categorical data analysis. It naturally develops the linear information models in contrast to the hierarchical log-linear models. A recent study by the authors (Journal of Data Science, Oct. 2006) clarified the difference in the data likelihood analysis between the two linear models. To supplement our previous study, a prototype of linear information models and a basic scheme of model selection are formulated. Empirical study with high-way contingency tables is employed to illustrate the validation of information models and the difference from the log-linear models.

How Many Factors? A Strategy for Identifying Latent Structure in Factor Analysis

❖ Mitchell Harris, Brigham Young University, 1001 SWKT, Psychology Department, Provo, UT 84602, *mitch_harris789@msn.com*; Michael K. Lauritzen, Brigham Young University; Landon Poppleton, Brigham Young University; Robert Bubb, Brigham Young University; Bruce Brown, Brigham Young University

Key Words: factor analysis, dimensionality, latent structure

Factor analysis has now been used for over a century. One of the persistent questions in using factor analysis has been the problem of determining when to quit factoring–that is, how does one know how many underlying factors should be used to capture the essence of the observed surface variables. A number of solutions to the number-of-factors problem have been proposed: the eigenvalues criterion, the skree plot method, etc. A Monte Carlo simulation method is used to create of a variety of surface variable structures from systematically varying underlying factor structures. This simulation has the advantage of evaluating the many and varied strategies for identifying the number of factors against the actual known number of factors in each simulation. A new strategy, based upon the central limit theorem, is shown to be most effective in identifying the actual number of factors.

Comparing Factor Loadings in Exploratory Factor Analysis: Comparison of Asymptotic and Randomization Tests

❖ Holmes Finch, Ball State University, Department of Educational Psychology, Muncie, IN 47306, *whfinch@bsu.edu*; Brian French, Purdue University

Key Words: multigroup factor analysis, invariance testing, exploratory factor analysis, randomization test

Multigroup Confirmatory Factor Analysis (MCFA) is useful for testing factorial invariance. MCFA requires model identification often by setting a factor loading equal across groups. A non-invariant referent loading can result in inaccurate tests and parameter estimation. A search procedure has been suggested for identifying an invariant referent loading, but is quite difficult to use. Exploratory Factor Analysis (EFA) also has been recommended for identifying invariant factor loadings through visually comparing loadings. However, this approach does not account for sampling variation. This study introduces two new tests for comparing loadings across groups: (1) an asymptotic test statistic using Maximum Likelihood EFA SE's, and (2) a randomization test statistic. The paper presents both methods, compares their performance in a simulation study, and provides an applied example.

Extensions to Latent Cluster Models for Social Networks

❖ Pavel Krivitsky, University of Washington, CSSS, C14 Padelford Box 354320, Seattle, WA 98195-4320, *pavel@stat.washington.edu*; Mark S. Handcock, University of Washington; Adrian E. Raftery, University of Washington; Peter D. Hoff, University of Washington

Key Words: social networks, latent variable models, model-based clustering, Bayesian inference

Latent space models for social networks postulate the existence of a latent ``social space', where the probability of a relation between entities depends on their relative positions within this space. The latent cluster model for social networks models groups of entities as model-based clusters on the latent space, and Bayesian fitting of this model via MCMC allows the latent space position estimation to borrow strength from the cluster process. We refine and extend this model in a number of ways, including reparametrizing to improve interpretability and reduce computational cost, modeling inhomogeneity through actor-level random effects, and generalizing the model to non-binary data. We demonstrate applications of this family of models to several datasets.

Estimating Personal Network Size

❖ Tyler H. McCormick, University of Connecticut, 207B Wu, Storrs, CT 06269, *tmccormick@stat.uconn.edu*; Tian Zheng, Columbia University; Matthew J. Salganik, Columbia University

Key Words: Social networks, Network size, Sampling, Overdispersion

Social network size provides information about the structure of relationships between individuals. This article presents a means of estimating network size and demonstrates how to design studies that estimate network size efficiently. We propose an estimation procedure based on count data reflecting the number of individuals with a particular first name known by the respondent. Estimation is done using a multilevel overdispersed Poisson regression model that reduces bias by accounting for the relationship between a respondent's age and the popularity of names over time. Mixing coefficients, which represent the proportion of a respondent's network that is of a particular age, are also estimated. We then derive simple estimates which, with appropriate age profiles, are self-weighting. Suggestions are given for survey designs that facilitate efficient estimation of network size.

391 Statistical Consulting in Games and Sports Organizations ●

Section on Statistics in Sports, Section on Statistical Consulting

Wednesday, August 1, 10:30 am–12:20 pm

Prediction of Hitting Performance in Baseball

❖ Shane T. Jensen, University of Pennsylvania, The Wharton School, 463 Huntsman Hall, 3730 Walnut Street, Department of Statistics, Philadelphia, PA 19104, *stjensen@wharton.upenn.edu*

Key Words: Baseball, Bayesian modeling, Prediction

There is substantial public and private interest in the projection of future performance in sports. We have developed a statistical model for the prediction of hitting performance for major league baseball players. The Bayesian paradigm provides a principled means for balancing past performance with crucial covariates, such as player age and position. We compare the performance of our model to current sabermetric methods on the 2005 season, and discuss both successes and limitations. We also discuss our procedure in the context of past investigations into shrinkage estimation of baseball performance.

It's All Fair Game: Consulting for an Online Competitive Entertainment Company

❖ Mark E. Glickman, Boston University, EN Rogers Mem Hosp (152), Bldg 70, 200 Springs Road, School of Public Health, Bedford, MA 01730, *mg@bu.edu*

Key Words: Handicapping, games, rating system, gambling

In the last 5+ years, there has been an explosion of web sites hosting skill-based games, such as traditional games like chess and poker, arcade-style games, and more. Some gaming sites induce players to compete by adding a gambling element; a player registers a credit card on the gaming site and then competes for cash and prizes in their favorite games. To ensure fairness and attract new players, the games need to be handicapped appropriately so that less-skilled players have a good chance at winning prize money. This talk discusses work performed by the author in conjunction with an online gaming company to develop a handicapping system for solitaire games. The work involved developing a statistical algorithm to produce a game score that the player would need to outperform in order to win a prize. The algorithm is fit and applied to game outcomes data from the gaming web site.

Academic Sports Radio: Construction of Interpretable and Valid Metrics for Sports Consulting

❖ C. Shane Reese, Brigham Young University, Department of Statistics, 230 TMCB, Provo, UT 84602, *reese@stat.byu.edu*

Key Words: sports, consulting

Sports is a data-rich, information-poor area of application. The distinction between data and information is explored through various interactions with the USA Olympic Committee, a National Football League team, and various local media outlets. In each of these three opportunities, the need for statisticians is obvious, as is the level of trust. We present three case studies, one from each of volleyball, football, and media interactions. Through three case studies, we argue the need for advisement on data collection, the creation of additional metrics, development of justifiable models, and the importance of proper interpretation and peer review.

392 Analyses and Applications in High-Throughput Screening Assays ● ✪

Biopharmaceutical Section, WNAR

Wednesday, August 1, 10:30 am–12:20 pm

High-Throughput Screening (HTS) for Lead and Target Identification

❖ Marc Ferrer, Merck & Co., Inc., 502 Louise Lane, North Wales, PA 19454, *marc_ferreralegre@merck.com*

Key Words: HTS, lead compounds, siRNA, drug discovery

High Throughput Screening (HTS) is now used to test large collections of compounds to identify leads for drug discovery. It is also used to test genome-wide collections of siRNA to probe biological processes to discover new therapeutically relevant targets and pathways. This talk will explain how the process of HTS fully automated for the rapid and cost effective screening.

Testing for Excess over Highest Single Agent with an Application to High-Throughput Screening

❖ John Peterson, GlaxoSmithKline, 709 Swedeland Road, UW281A, King of Prussia, PA 19406-0939, *john.peterson@gsk.com*

Key Words: Adjusted p-values, bootstrapping, combination drug study, factorial design, trend test

Combination drug therapy offers much promise for discovering pharmaceutical treatments that are efficacious and safe. A key efficacy criterion for a combination of two compounds is that the combination is superior to both of its component compounds used alone. This article proposes a simultaneous testing procedure, based upon step-down trend tests, that identifies dose combinations for pairs of compounds that produce efficacy results with excess over highest single agent (i.e., the combination is superior to both of the component compounds). This testing procedure is applied to data from experiments for a pilot high-throughput screening study for pairs of compounds evaluated at nine dose levels using 9x9 factorial experiments. This procedure is easily automated and can be computed using the SASÆ MULTTEST procedure.

Simple Statistical Parameters but Quick Solutions in High-Throughput Screening Assays

❖ Xiaohua (Douglas) Zhang, Merck & Co., Inc., Biometric Research, WP53B-120, West Point, PA 19486, *xiaohua_zhang@merck.com*

Key Words: High throughput screening, Strictly standardized mean difference, Z-factor, RNA interference

High throughput screening (HTS) plays a central role in modern drug discovery, allowing the rapid screening of diverse chemical compounds (i.e., small molecule) or short interfering RNAs (siRNAs) to identify active compounds or effective siRNAs (“hits”). For example, small molecule HTS allows for screening of more than 100,000 compounds a day per screen. Thus, HTS experiments generate a huge amount of data. In order to use these data to identify the most active compounds or effective siRNAs tested, it is critical to adopt and develop appropriate statistical methods for quality control (QC) and hit selection. Here, I present some simple statistical parameters and explore their uses in both quality control and hit selection in RNAi HTS assays.

Applications of RNAi High-Throughput Screening in Drug Discovery

❖ Amy S. Espeseth, Merck & Co., Inc., 770 Sumneytown Pike, West Point, PA 18914, *Amy_Espeseth@merck.com*

Key Words: siRNA, High Throughput Screening, cell-based assays

The discovery of RNA interference (RNAi), which allows the targeted degradation of mRNAs using small interfering RNA (siRNA), has revolutionized the identification of drug discovery targets. Following the development of a cell-based assay for human disease, cellular arrays can be treated with siRNA libraries that target thousands of genes. Downregulation of genes associated with the disease state affects the assay output, similar to what would be expected from a small molecule inhibitor. Thus effective siRNAs that are identified in siRNA screens point to potential novel targets for drug discovery programs. Although the number of data points generated by these screens is small compared to a typical small molecule HTS, the number of variables introduced by the use of siRNA typically increases the assay variance, complicating the use of typical QA and hit selection statistical methods.

393 Statistical Issues and Novel Models for Quantifying the Effects of Continuous Traits on Mortality: Examples from Obesity ● ✪

ENAR, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Biometrics Section, WNAR

Wednesday, August 1, 10:30 am–12:20 pm

Methods for Correcting Regression Dilution and Reverse Causation Biases in Regression-Model Survival Analyses

❖ James Greenberg, Brooklyn College, 2900 Bedford Avenue, Brooklyn, NY 11210, *jamesg@brooklyn.cuny.edu*

Key Words: Within-person variation, measurement error, regression dilution, reverse causation, body mass index, blood pressure

Three methods were tested: 1) correcting for regression dilution by using the average of several replicates as predictor; 2) correcting for reverse causation by excluding ever smokers and also excluding participants with age>60 yrs and evidence of secular decrease in predictor; and 3) correcting for both biases by using the average of several replicates as predictor and the baseline observation as a covariate. These methods were tested using data from two prospective cohorts. Two predictors were tested, blood pressure and body mass index. All three methods were found to cause substantial increases in hazard ratios at elevated levels of the risk factor. These methods therefore appear to be effective in correcting for some of the effects of regression dilution and reverse causation. Further research is need to uncover their modus operandi, and to define when their use is appropriate.

Intentional Weight-Loss Effects on Mortality Rate Modeled as a Latent Variable Problem

❖ Gary L. Gadbury, University of Missouri-Rolla, Mathematics and Statistics, 202 Rolla Building, Rolla, MO 65409, *gadburyg@umr.edu*

Key Words: causation, nonestimable, obesity, potential outcomes

Obesity is associated with increased mortality rate and short-term weight loss improves risk factors for mortality. It has not been convincingly shown that weight loss among obese persons results in reduced mortality rate.

Studies have pointed out that weight loss is often a sign of illness and intentional weight loss (IWL) should be separated from unintentional weight loss (UWL). Coffey et al. (Stats in Med, 2005) considered the problem of disentangling the effects of IWL on mortality from UWL. They showed that the effect of IWL on mortality is a nonestimable parameter. A potential outcomes framework is presented to illuminate required assumptions and for defining estimable bounds for the nonestimable parameter. The framework is extended to the scenario where subjects change their intention to lose weight over two time periods. A Bayesian approach to this problem is also discussed.

In Search of the Law of Mortality

❖S. Jay Olshansky, University of Illinois at Chicago, 400 Springside Lance, Buffalo Grove, IL 60037, *sjayo@uic.edu*

Key Words: mortality, aging, longevity, evolution

The duration of life of most sexually reproducing species, including humans, has been shown to be calibrated to an attribute of the life history that has nothing at all to do with aging—reproduction. The characteristic J-shaped curve of mortality is also a common attribute of many forms of life, and this too is calibrated to the onset and length of the reproductive window. For more than 180 years scientists speculated that the common age pattern of death represented a "law of mortality", but for a variety of reasons no one was able to demonstrate its existence. In a test of this hypothesis, we have demonstrated that the law of mortality exists, and that there are evolutionary explanations for its presence. Experimental verification of the law of mortality will be presented using unique sources of mortality data on humans, dogs, and mice.

Accounting for Unobserved Confounders via Data Augmentation

❖David B. Allison, The University of Alabama at Birmingham, Dept. of Biostatistics, Ryals Public Health Bldg, 420C, Birmingham, AL 35294, *dallison@uab.edu*

Key Words: causal effects, adiposity, mortality rate, putative confounding, weight loss

A recurrently identified challenge in the analysis of data seeking to estimate the causal effects of adiposity on mortality rate is the putative confounding by unobserved variables that both predispose to weight loss and to earlier death. To correct for this, investigators have utilized a variety of ad hoc procedures that typically entail eliminating subsets of subjects in the hopes that the analyses of remaining subjects will yield unbiased estimates. Unfortunately, such ad hoc procedures enjoy little statistical foundation and have not been validated. I will propose an alternative approach in which information on putative but unobserved confounders is taken from other studies and incorporated into analyses of the dataset in hand using techniques adapted from multiple imputation. The approach will be illustrated with both real and simulated data.

394 Understanding and Measurement of Human Capital ● ✪

Social Statistics Section, Section on Survey Research Methods

Wednesday, August 1, 10:30 am–12:20 pm

The Whole Is Greater than the Sum of Its Parts: The U.S. Census Bureau Human Capital Management Information System

❖Juanita Tamayo Lott, U.S. Census Bureau, 4700 Silver Hill Road, Washington, DC 20233-1400, *juanita.t.lott@census.gov*; Tyra D. Smith, U.S. Census Bureau; Mark Holdrege, U.S. Census Bureau

Key Words: human capital, federal standards, statistical standards, management information systems

Human capital is a phenomenon that can be examined within a statistical framework. This paper focuses on the U.S. Census Bureau Human Capital Management Information System (HCMIS) as a fundamental managerial and analytical tool to monitor ongoing human capital and to anticipate changes that may require review or adjustment to human capital. We address four areas: principles and standards for federal and statistical systems, the HCMIS as a management and analytical tool, understanding and measurement of human capital based on use of HCMIS in 2003–2007, and a human capital vision for the 2008–2013 period, which leverages talent and technology.

Measuring a Blended 21st-Century Work Force: Lessons from NASA

❖Stephen E. Chesley, National Aeronautics and Space Administration, 300 E Street SW, Mailstop 4U70 A, Washington, DC 20546-0001, *Stephen.Chesley@nasa.gov*; Craig Conlin, National Aeronautics and Space Administration

Key Words: NASA, Workforce, Government, Measuring, Public-sector, Private-sector

In the 21st Century, government agencies have realized that their work force consists of more than just full-time permanent civil service employees. Their workforce also comprises part-time employees, temporary/term/intermittent workers, private sector contractors, grantees, scholars from academia, and even employees that are employed by other government agencies. This paper explains some of the efforts that NASA is making to fully understand its work force. This paper begins by explaining some of the methods that NASA has used in the past to understand and to measure its work force. Next, this paper will look at NASA's current work force; how NASA determines what its work force is, how it is managing this work force, and how it plans on managing its work force in the future. The concluding section will discuss possible ways organizations can better measure a blended work force.

Accurately Measuring a Creative Economy

❖Steven W. Pedigo, Greater Washington Initiative, 1725 Eye Street, NW, Suite 200, Washington, DC 20006, *stevenpedigo@bot.org*

Key Words: workforce, human capital, occupational, talent, Greater Washington, creative

Over the past 10 years, the creative sector has expanded in size and significance with remarkable speed. More than ever, mature regional economies rely on their creative cores for added productivity, value, and employment. Although economic development practitioners and researchers recognize a strong creative sector is a cornerstone of a healthy and mature economy, the relatively recent emergence of the sector has left suitable, accurate measures in a nascent stage. An occupational analysis offers the most prudent measure of the creative sector, showcasing the individual as the driver of the sector. Using the greater Washington, DC, USA, region as a case study, this presentation addresses the methods and policy implications of a creative occupational analysis.

Water, Water Everywhere and Not a Drop to Drink: Higher Education Access and Success Among America's Growing Population of Under-Represented Students of Color

❖ Walter R. Allen, University of California, Los Angeles, 3101A Moore Hall, Grad Sch Education, Los Angeles, CA 90095, *wallen@ucla.edu*; Derek Mitchell, Stupski Foundation; Gloria Gonzalez, University of California, Los Angeles

Key Words: Human Capital, Achievement, Race, Gender, Web Study

Latinos and Blacks are a large, growing share of California's K–12 students, yet both groups are under- represented and declining on college campuses. Do teacher biases influence student assessments and decisions to nominate students for undergraduate university admissions? Teachers are recruited from a large, urban school district. Using web-based simulation we examine whether student racial, ethnic or gender background (displayed in photographs) influence teacher decisions to nominate students for university admission. We expect that teacher evaluations and nominations will be biased to favor whites and Asians and against Latinos and Blacks. If race "colors" how teachers read and evaluate objective academic information, we discover an important source of persistent racial inequality in achievement. We recommend professional development activities to improve learning outcomes.

395 Mighty Nonparametrics ●✪

Noether Award Committee, General Methodology

Wednesday, August 1, 10:30 am–12:20 pm

Nonparametric Modeling for Classification of High-Dimensional Data and Multiple Hypothesis Testing

❖ Peter Hall, The University of Melbourne, Department of Mathematics and Statistics, Melbourne, 3010 Australia, *peter.hall@maths.anu.edu.au*

Key Words: TBD

Problems involving classification of high-dimensional data, and `highly multiple' hypothesis testing, arise frequently in the analysis of genetic data and complex signals. Their theoretical elucidation raises challenges, however. We address this by interpreting small samples of high-dimensional data. Depending to some extent on how erratic the time-series are, important features of classifiers, or of multiple hypothesis testing procedures, can be accessed by exploring properties of time-series models. For example, it can be shown that, in the context of multiple hypothesis testing, the assumption of independence is much less of an issue in high-dimensional settings than in conventional, low-dimensional ones. Similar arguments can be employed to explore other aspects of the analysis of high-dimensional data.

Nonparametric Hierarchical Models

❖Ciprian M. Crainiceanu, Johns Hopkins University, 615 N. Wolfe St. E3636, Department of Biostatistics, Baltimore, MD 21205, *ccrainic@jhsph.edu*

Key Words: signal extraction, hierarchical smoothing, penalized splines

We develop the nonparametric hierarchical inferential framework with focus on model estimation and testing. Two complementary methodologies are discussed and implemented for scientific and geometric signal extraction. We identify, describe and solve a new measurement error problem generated by the hierarchical smoothing process when signals are used in second level analyses. Our research was motivated by a large (6,400 subjects) longitudinal study of electro encephalographic (EEG) signals during sleep where the natural hierarchy is patient/group or cluster/visit. The design of our methods and use of low rank penalized splines make our computations tractable.

396 Statistical Models and Methods for Microarray Data

IMS, Biometrics Section, Biopharmaceutical Section, WNAR

Wednesday, August 1, 10:30 am–12:20 pm

Inferences on the Proportion of Non-Null Effects in Large-Scale Multiple Comparisons

❖ Jiashun Jin, Purdue University, 250 N. University St., Department of Statistics, West Lafayette, IN 47907, *jinj@stat.purdue.edu*

Key Words: Multiple testing, Proportion of non-null effects, Oracle, Uniform consistency

The immediate need for fast and effective data mining for massive data gives rise to a recent new field in statistics: large-scale multiple comparisons, where thousands or even millions of hypotheses are testing simultaneously. In this talk, we consider a problem of particular interest in this field: estimating the proportion of non-null effects. We report some recent progress on estimating the proportion, ranging from the non-sparse regime to the very sparse regime. In particular, we report a method for constructing oracle equivalence of the proportion, which equals to the proportion for any underlying parameters. The oracle naturally yields real estimators, which are shown to be uniformly consistent for the proportion over a wide class of situations. This talk is based joint works with (alphabetically) Tony Cai, David Donoho, Mark Low, Jie Peng, and Pei Wang.

Epigenomics: Some Statistical Applications

❖ Rafael Irizarry, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD , *rafa@jhu.edu*

Key Words: epigenetics, microarrays

The use of modern statistical methodology can substantially improve accuracy and precision of gene expression microarray measurements. Microarrays are now being used to measure diverse high genomic endpoints including epigentic marks. In this talk we will describe some of the issues arising in Epigenetic applications and how some of these are similar to expression applications and other are not. For example, we will demonstrate how feature intensities seem to follow additive background/multiplicative error models similar to those observed for expression data and also how similar probe sequence effect are seen. As for the differences, we will show how some of the assumption made in expression studies to normalize arrays (a step needed to enable across array comparisons) are no longer valid. We will briefly describe some of our current ideas to overcome the new issues.

Simultaneous Subset Selection via Rate-Distortion Theory with Applications to Clustering and Significance Analysis of Gene Expression Data

❖ Rebecka Jornsten, Rutgers University, Department of Statistics, 501 Hill Center, Busch campus, Piscataway, NJ 07030, *rebecka@stat.rutgers.edu*

Key Words: model selection, clustering, FDR, significance analysis

We present a novel model selection methodology, based on the well-known principle of bit-allocation in rate-distortion theory. The method turns the combinatorial subset selection problem, across clusters and variables, into a simple line search. Moreover, in the special case where data objects form their own clusters, the methodology generalizes to subset selection in a multiple testing framework. The selection method is applied to a time course gene expression dataset of two proliferating stem cell lines. Our analysis provides sparse and easy-to-interpret gene cluster models. In addition, by incorporating subset selection into testing, the power of detection is significantly increased. The method provides sparse model representations for each gene. Thus, we find that the significant list is dominated by main cell line and time effects, and a few informative interaction models.

397 Causal Methods: Advances and Applications in Health ● ✪

Section on Health Policy Statistics, Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences, WNAR

Wednesday, August 1, 10:30 am–12:20 pm

Gangs and Teen Violence: A Matched Analysis Using Trajectory Groups and Propensity Scores

❖Amelia Haviland, RAND Corporation, 4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665, *haviland@rand.org*; Daniel Nagin, Carnegie Mellon University; Paul Rosenbaum, University of Pennsylvania

Key Words: mixture model, propensity score, optimal matching, observational study, trajectory group

Using data from the MontrÈal Longitudinal-Experimental Study of Boys, the effects on subsequent violence of joining a gang at age 14 are studied controlling for characteristics of boys prior to age 14. The boys are divided into trajectory groups based on violence from ages 11 to 13, and within trajectory groups, joiners are optimally matched to a variable number of controls using propensity scores, Mahalanobis distances, and a combinatorial optimization algorithm. The trajectory groups define meaningful subpopulations where effects may be different, while propensity scores and optimal matching tend to balance twelve baseline covariates. By using between 1 and 7 controls for each joiner, greater efficiency is obtained than by pair matching, with greater bias reduction than is available by matching in a fixed ratio.

Causal Inference of Nonrandomized Treatment Effects

❖Heejung Bang, Cornell University, 411 E. 69th St., New York, NY 10021, *heb2013@med.cornell.edu*; Yolanda Barrón, Cornell University

Key Words: non-randomized treatment, causal inference, marginal structural model, double-robustness

Randomized clinical trials are the gold standard used to evaluate the "causal" effects of treatments in biomedical research. However, it is not feasible to conduct a clinical trial to answer every clinical question. Nonrandomized treatments can not serve as a valid alternative, but often appear in various contexts within observational studies or even as secondary interventions within randomized clinical trials. We use the causal models (e.g., marginal structural model) in cross-sectional and longitudinal settings for the analysis of a non-randomized intervention and further extend the models for doubly robust estimation. The proposed methods are applied to study the effect of non-randomized antidepressant use by depression suffering patients enrolled in a cardiovascular clinical trial, on myocardial infarction and death.

Dealing with Control Group Drop-Out: Using Historical Patient Information to Supplement a Randomized Trial

❖Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, 8th Floor, Baltimore, MD 21205, *estuart@jhsph.edu*; Donald B. Rubin, Harvard University; Samantha Cook, Google

Key Words: propensity score, causal inference, matching methods, missing data, Bayesian methods, patient registries

The goal of a clinical trial is often to compare a new treatment with currently available treatments. However, in some trials, some control group members may receive the new treatment and thus can no longer be considered controls. We describe a method of utilizing information from a large collection of "traditionally treated" patients (e.g., available through patient registers) to impute the control patients' outcomes as if they had stayed on the traditional treatment. This is done by modeling trends in the disease using the individuals in the patient register who look most similar to the study patients, selected through propensity score matching. This talk focuses on two particular challenges to the matching: missing covariate data and the need to define baseline for patients in the register. The methods are illustrated using data from a clinical trial for a rare disease.

Propensity Scores and Multiple Treatments: An Application to Beneficiary Evaluations from the Medicare CAHPS Surveys

❖Marc N. Elliott, RAND Corporation, 1776 Main St., Santa Monica, CA 90407-2138, *elliott@rand.org*; Amelia Haviland, RAND Corporation

Key Words: Multinomial, Prescription Drugs, Insurance, Consumer Evaluation of Healthcare, Observational Data, Propensity Score

The Centers for Medicare and Medicaid Services provides comparisons of beneficiary experiences with original and managed care Medicare. In this paper, we use the 2004 Medicare CAHPS surveys to compare beneficiary perceptions of care across five different types of Medicare health coverage that include original Medicare and managed care plans with and without prescription drug coverage and other supplementary coverage. Enrollees are not randomly assigned to types of health coverage, and there are well documented differences between the enrollees across these five options. This paper compares estimates of beneficiary experience using regression methods to 'control' for differences in enrollee characteristics to estimates obtained from a propensity score analysis for multiple groups. We also use related models to demonstrate that the "optimal" coverage varies by enrollee characteristics.

398 New Developments on Quantile Regression ✪

IMS

Wednesday, August 1, 10:30 am–12:20 pm

Bayesian Model-Based Approaches to Semiparametric and Nonparametric Quantile Regression

❖Athanasios Kottas, University of California, Santa Cruz, Department of Applied Math and Statistics, School of Engineering, 1156 High Street, Santa Cruz, CA 95064, *thanos@ams.ucsc.edu*; Milovan Krnjajic, Lawrence Livermore National Laboratory; Matthew Taddy, University of California, Santa Cruz

Key Words: Censored quantile regression, Dirichlet process mixture models, Markov chain Monte Carlo, Multivariate normal mixtures, Scale uniform mixtures

We present two approaches to Bayesian quantile regression, using Dirichlet process (DP) mixture models. We first consider a semiparametric additive formulation, with the regression function modeled parametrically and with nonparametric priors for the error distribution. The prior models include dependent DPs resulting in error distributions that can change nonparametrically with the covariates. The second line of research develops a flexible fully nonparametric approach to inference for any set of quantiles of the response distribution. Under this approach, the joint distribution of the response and the covariates is modeled with a DP mixture, with posterior inference for different quantile curves emerging through the conditional distribution of the response given the covariates. Inference (possibly, under censoring) is implemented using posterior simulation methods for DP mixtures.

Some Recent Advances in Computing Regularized Quantile Regression Models

❖ Ji Zhu, University of Michigan, 439 West Hall, 1085 South University Ave., Ann Arbor, MI 48109-1107, *jizhu@umich.edu*

Key Words: Degrees of freedom, Linear programming, Quadratic programming, Quantile regression, Regularization, RKHS

Classical regression methods have mainly focused on estimating conditional mean functions. However, in recent years, quantile regression has emerged as a comprehensive approach to the statistical analysis of response models. In this talk we consider two regularized quantile regression models, the L1-norm (LASSO) regularized quantile regression and quantile regression in reproducing kernel Hilbert spaces. We propose efficient algorithms that compute entire solution paths of the two models. We also derive simple formulas for the effective dimensions of the two models, which allow convenient selection of the regularization parameters.

Learnability and Robustness of Nonparametric Quantile Regression

❖ Andreas Christmann, Free University of Brussels VUB, Department of Mathematics, Pleinlaan 2, Brussels, 1050 Belgium, *Andreas.Christmann@vub.ac.be*; Ingo Steinwart, Los Alamos National Laboratory

Key Words: quantile regression, empirical risk minimization, consistency, kernel, non-parametric

Quantile regression is used in many areas of applied research and business. Examples are actuarial, financial or biometrical applications. We show that a non-parametric generalization of quantile regression based on kernels proposed by Takeuchi et al. (2006) shares with support vector machines the property of consistency to the smallest possible risk (i.e., to the Bayes risk). Further we use this consistency to prove that the non-parametric generalization approximates the conditional quantile function which gives the mathematical justification that kernel based quantile regression is able to learn. Some results concerning robustness properties of such regression estimators are also given.

Conditional Quantile Estimation for GARCH Models

❖ Zhijie Xiao, Boston College, Dept of Economics, Chestnut Hill, MA 02467, *xiaoz@bc.edu*

Key Words: Quantile Regression, GARCH, Nonlinear

The purpose of this paper is to propose a robust, flexible approach to estimating conditional quantiles in time series with GARCH structure. Conditional quantiles is an essential ingredient in various risk measures, and the GARCH process has proven to be highly successful in modeling financial return data. However, quantile regression estimation of GARCH models is highly nonlinear, and can not be directly estimated by traditional recursive methods. We propose two new methods of estimating quantiles of GARCH models. The first method is based on minimum-distance estimation from a first stage nonlinear quantile regression. The second method is based on a preliminary sieve quantile regression. Asymptotic properties of both methods are investigated.

399 Wald Lecture II

IMS, General Methodology, Section on Bayesian Statistical Science
Wednesday, August 1, 10:30 am–12:20 pm

Model Selection: Multiplicities and Approximations

❖ James Berger, Duke University, Inst of Statistics and Decision Science, Durham, NC 27708-0251, *berger@stat.duke.edu*

Key Words: BIC, multiplicity, reproducability, multiple testing, asympotics, marginal likelihood

Two of the biggest hurdles in dealing with model uncertainty are dealing with multiplicity and computation. Issues of multiplicity in testing are increasingly being encountered in practice, and failure to properly adjust for multiplicities is being blamed for the apparently increasing lack of reproducibility in science. The first part of the talk will discuss different types of multiplicities that are being encountered and methods for handling them. A major computational hurdle for Bayesians is computation of model likelihoods. Because of the difficulty of this computation, BIC is often used as an approximation. Unfortunately, BIC has a number of problems. The second part of the talk will cover new approximations that show considerable promise in significantly improving on BIC, even potentially applying to situations where the model size grows with the sample size.

400 Frailty Models for Survival Analysis ●

Biometrics Section, ENAR, Section on Risk Analysis
Wednesday, August 1, 10:30 am–12:20 pm

Semiparametric Approaches for the Analysis of Multilevel Failure Time Data

❖ Joanna H. Shih, National Cancer Institute, 6130 Executive Blvd, EPN room 8132, Bethesda, MD 20892-7434, *jshih@mail.nih.gov*; Shou-En Lu, University of Medicine and Dentistry of New Jersey

Key Words: multi-level clustering, nested frailties, Monte-Carlo EM, Within-cluster resampling

Multilevel clustered failure time data arise when clustering of data occurs at more than one level. Often it is of interest to make inference on the association of failure times at each level of clustering. Two modeling approaches (i.e., marginal and conditional) are considered for analyzing this type of data. For the marginal approach, we consider a class of multivariate survival models parameterized by marginal distributions and accounting for hierarchical structure of clustering through copula functions. This class of models emphasizes on a population-averaged interpretation for the covariate effects on the marginal hazard. For the conditional approach, we reparameterize the above class of models such that they are represented as nested random effects proportional hazard models.

Frailty Models for Survival Data with Informative Censoring

❖ Xuelin Huang, The University of Texas M.D. Anderson Cancer Center, Department of Biostatistics and Applied Mathematics, 1515 Holcombe BLVD, Unit 447, Houston, TX 77030, *xlhuang@mdanderson.org*; Robert

Wolfe, University of Michigan; Lei Liu, University of Virginia

Key Words: Clustered data, Competing risks, Dependent censoring, EM algorithm, Recurrent event, Survival analysis

A new class of frailty models is used to account for not only the correlation between failure times, but also the correlation between failure and censoring. These models have wide applications in the analyses for clustered survival data, recurrent event data and others. The proposed models allow multiple types of censoring. Censoring by some causes can be analyzed as informative, while censoring by other causes are assumed to be non-informative. To fit the model, the EM algorithm is used with Markov Chain Monte Carlo simulations in the E-steps. The EM algorithm in the setting of semi-parametric frailty models involves some complications, which will be illustrated and clarified in detail. Simulation studies show that the proposed models and the fitting method work well.

Goodness-of-Fit for Frailty Models

❖John Klein, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, *klein@mcw.edu*

Key Words: Copula Models, Frailty Models

Random effects or frailty models have become very useful in modeling dependence between individual survival times or between survival times in multistate modeling. A number of models have been proposed for the frailties in the literature but it is difficult in practice to pick between these models. In this presentation we present a general method to asses the goodness of fit of the choice of the frailty distribution. This technique, which is essentially a goodness of fit test for the copula, is based on a comparison between a distribution free estimator of the copula and the postulated copula. The method is available for any copula and does not involve the marginal survival functions.

401 Remote Sensing in Environmental Statistics ● ✪

Section on Statistics and the Environment, Section on Statisticians in Defense and National Security

Wednesday, August 1, 10:30 am–12:20 pm

Markov Random Fields, Remote Sensing, and Predicting Crop Yields

❖Stephan Sain, National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, CO 80305, *ssain@ucar.edu*

Key Words: Spatial/temporal models, Multivariate models, Maximum likelihood

Predicting crop yields is a problem of vital interest to researchers and policy makers. Remote sensing data can be of great use in this problem as it is a useful integration of weather, farming practice, etc. In this talk, we will incorporate remotely sensed data into a multivariate, spatial-temporal Markov random field model to simultaneously predict yields of corn and wheat. Details of the model will be covered as well as a computationally efficient approach to parameter estimation. Results will also be discussed.

Nonstationary Covariance Models for TOMS Ozone Data

❖Mikyoung Jun, Texas A&M University, *mjun@stat.tamu.edu*

Key Words: spherical process, Nonstationary spatial covariance function, TOMS ozone data

TOMS ozone data show complicated nonstationary covariance structure. In particular, there is a clear dependency of the covariance structure on latitude and Northern and Southern hemisphere may show different patterns of covariance structure as well. We model TOMS ozone data on the globe through a rich and flexible nonstationary covariance models, which is an extension of the general model framework in Jun and Stein (2004). Empirical measures of coefficients of differential operators for the models in Jun and Stein (2004) will be given. Optimal interpolation results based on these models and computational issues in calculating the covariance matrices will be discussed.

Understanding Large-Scale Structure in Massive Remote-Sensing Datasets

❖Amy Braverman, Jet Propulsion Laboratory, California Institute of Technology, Mail Stop 126-347, 4800 Oak Grove Drive, Pasadena, CA 91106, *Amy.Braverman@jpl.nasa.gov*

Key Words: remote sensing, massive data sets, data compression, hypothesis testing, atmospheric science

NASA's Atmospheric Infrared Sounder (AIRS) mission has been collecting remote-sensing data about the vertical structure of Earth's atmosphere since AIRS was launched in mid 2002. The AIRS team has implemented the Level 3 Quantization data product to provide users with monthly summaries of the joint distributions of measurements at 11 altitudes of atmospheric temperature, water vapor content, and cloud fraction. In each month, each five-degree spatial grid cell contains a multivariate distribution estimate. In this talk we discuss hypothesis testing procedures for characterizing trends, identifying outliers, and distinguishing between atmospheric regimes based on distributional comparisons.

Space-Time Modeling of Biomass Burning and Regional Aerosols in Southeast Asia

❖Catherine A. Calder, The Ohio State University, 1958 Neil Ave., Department of Statistics, Columbus, OH 43210, *calder@stat.ohio-state.edu*; Tao Shi, The Ohio State University

Key Words: Atmospheric Science, Bayesian Modeling, Remote Sensing, Spatial Statistics

Scientists and policy makers have become increasingly concerned about the implications of the consistent brown haze covering Southeast Asia in terms of human health and climate change. The emergence of this haze is due to increased atmospheric concentrations of carbonaceous aerosols, which are generated by anthropogenic activities including both slash-and-burn agriculture and fossil fuel combustion. Our research focuses on determining the relative contribution of these two types of emissions to the total aerosol burden over the region. We propose a space-time model for regional carbonaceous aerosol composition and concentration, given atmospheric circulation processes and observed fire occurrence. Our model synthesizes a variety of types of data including remote sensing imagery, output from atmospheric transport models, and estimates of biomass emissions for various vegetation types.

402 Applications in Computational Biostatistics

Section on Statistical Computing, ENAR

Wednesday, August 1, 10:30 am–12:20 pm

Application of the State-Space Modeling Paradigm in Neuroscience Data Analysis

❖Emery N. Brown, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, *brown@neurostat.mgh.harvard.edu*

The state-space paradigm has been widely used to analyze a broad range of stochastic dynamical systems problems in engineering, computer science, statistics and the social sciences. We have used the paradigm to construct signal processing algorithms to analyze several neural systems. In this talk, we will review our work on the use of the state-space modeling paradigm to study three problems in neuroscience data analysis: tracking on a millisecond time-scale the dynamics of the spatial receptive fields of rat hippocampal neurons during learning; decoding how ensembles of neurons in the rat hippocampus maintain a dynamic representation of the animal's position in its environment; and devising a dynamic solution to the source localization problem for magnetoencephalography.

The Use of Distances in Surveillance

❖ Marcello Pagano, Harvard University, School of Public Health, Boston, MB 02115-6023, *pagano@hsph.harvard.edu*

Key Words: distances, surveillance, high-dimensions

The distribution of distances between individuals is the focus of this talk. We show examples of where the study of such distances is useful in the detection of patterns and relationships. These may be studied even when the distribution of individuals is not known or hypothesized, a characteristic that becomes more important as the dimensionality of the data space increases. We show how to test hypotheses about these distributions, how to estimate prevalence in multidimensional spaces, and how to perform regressions.

Assessing Interventions Related to HIV Incidents Under the Influence of Ethanol

Yasmin H. Said, George Mason University; ❖ Edward Wegman, George Mason University, 4400 University Drive, Fairfax, VA 22030-4422, *ewegman@gmu.edu*

Key Words: social networks, alcohol abuse, disease mechanism

Alcohol abuse leads to serious individual and societal outcomes. Among these, we identify irresponsible behavioral outcomes notably risky sexual contacts and other violence related outcomes. Risky sexual contacts can lead to infections with HIV and other STDs. Social, structural, and environmental factors are major influences on HIV-related behaviors. Our work provides a policy tool for evaluation of interventions using a dynamic agent-based simulation based on estimates of conditional probabilities. Alcohol abusers are embedded in a social network that is reminiscent of an ecology system. We formulate a model of this network to explore interventions that reduce the overall probability of negative outcomes. Historically, modeling attempts have focused on specific negative outcomes. The unique feature of this work is that we explore the simultaneous reduction of all negative outcomes.

403 Advances in Statistical Quality Improvement ● ✪

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Wednesday, August 1, 10:30 am–12:20 pm

Bayesian Modeling of Accelerated Life Tests with Random Effects

Ramon V. Leon, University of Tennessee; ❖ Avery J. Ashby, BlueCross BlueShield of Tennessee, *Avery_Ashby@BCBST.com*

Key Words: Bayesian Life Tests, Markov chain Monte Carlo (MCMC), Random Effects, Credibility Interval, Prediction Interval, WinBUGS

We show how to use Bayesian methods to make inferences from an accelerated life test where the test units come from different groups (such as batches) and the group effect is random and significant both statistically and practically. Our approach can handle multiple random effects and several accelerating factors. We illustrate the method with an application concerning Space Shuttle pressure vessels wrapped in Kevlar 49 fibers where the fibers of each vessel comes from a single spool and the spool effect is random. We show how Bayesian analysis using Markov chain Monte Carlo (MCMC) methods is used to answer questions of interest in accelerated life tests with random effects which are not as easily answered with more traditional frequentist methods. Our approach is implemented in the freely available WinBUGS software so that readers can apply the method to their own data.

A Control Chart for the Coefficient of Variation

❖ Chang W. Kang, Hanyang University, 1271 Sa-1-dong, Sangnok-ku, Ansan, 426-791 South Korea, *cwkang57@hanyang.ac.kr*; Man S. Lee, LG Philips LCD; Young J. Seong, Gyenggi Research Institute; Douglas M. Hawkins, The University of Minnesota

Key Words: : Statistical process control, process variability, clinical chemistry

Monitoring variability is a vital part of modern statistical process control. The conventional Shewhart R and S charts address the setting where the in-control process readings have a constant variance. In some settings, however, it is the coefficient of variation, rather than the variance, that should be constant. For example this setting is common in clinical chemistry, and then conventional R and S charts cannot be used. This paper develops a chart, equivalent to the S chart, for monitoring the coefficient of variation using rational groups of observations.

An Optimal Filter Design Approach to Statistical Process Control

❖ Daniel W. Apley, Northwestern University, Dept. of Industrial Engineering, 2145 Sheridan Rd., Evanston, IL 60208, *apley@northwestern.edu*; Chang-Ho Chin, Kyung Hee University

Key Words: Statistical Process Control, Control Charts, Autocorrelation, Linear Filtering, Markov Chains

Many control charts can be viewed as charting the output of a linear filter applied to process data, with an alarm sounded when the filter output falls outside a set of control limits. We generalize this concept by considering a linear filter in its most general time-invariant form. We provide a strategy for optimizing the filter coefficients in order to minimize the out-of-control ARL, while constraining the in-control ARL to some desired value. The optimal linear filters exhibit a number of interesting characteristics, in particular when the process data are autocorrelated. In many situations, they also substantially outperform an optimally designed exponentially weighted moving average (EWMA) control chart.

The Rate of False Signals in X-bar Control Charts with Estimated Limits

❖ Diane P. Bischak, University of Calgary, Haskayne School of Business, 2500 University Dr NW, Calgary, AB T2N1N4 Canada, *diane.bischak@haskayne.ucalgary.ca*; Dan Trietsch, American University of Armenia

Key Words: Average run length, Diffidence, Control limits, Estimated parameters

The in-control statistical properties of X-bar charts have usually been studied from the perspective of the average run length until the first (false) signal, known as the in-control ARL. We argue that the ARL is a confusing concept when used with charts with estimated limits and that the rate of false signals (RFS), which focuses on the behavior of charts during extended use, is more intuitive. We use the RFS to illustrate graphically the dangers of using too few subgroups to estimate control limits. We also discuss diffidence charts, which make the inherent uncertainty concerning RFS observable to the practitioner and thus help the practitioner determine what is an acceptable number of subgroups for a given charting application.

404 2007 ASA Educational Ambassadorship

Committee on International Relations in Statistics

Wednesday, August 1, 10:30 am–12:20 pm

Challenges with Statistical Education in Developing Countries: A Case Study

❖Eshetu Wencheko, Alpha University College, P.O. Box 5771, Addis Ababa, Ethiopia, *ewencheko@yahoo.com*

Although the demand for skilled statisticians is growing in developing countries, measures taken to address the shortage have not been satisfactory. The recently launched ASA Educational Ambassadorship program is a positive step in raising awareness about the gap, and in providing much needed support to enhance statistical education in these countries. In this talk, we highlight some of the problems, and share the experience gained as an ambassador.

405 Estimating Nonresponse Bias in Light of the New OMB Standards and Guidelines for Statistical Surveys ● ✪

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Wednesday, August 1, 10:30 am–12:20 pm

Estimating Nonresponse Bias in Light of the New OMB Standards and Guidelines for Statistical Surveys

❖J. Michael Brick, Westat, 1650 Research Blvd., RE 494, Rockville, MD 20850, *mikebrick@westat.com*; ❖Robert M. Groves, University of Michigan, 426 Thompson Street, Survey Research Center, Ann Arbor, MI , *bgroves@isr.umich.edu*; ❖Brian Harris-Kojetin, Office of Management and Budget, 725 17th Street, NW, Room 10201, Washington, DC 20503, *Brian_A._Harris-Kojetin@omb.eop.gov*

Key Words: OMB Guidelines, Nonresponse Bias, Response rates, Bias estimates, Bias anaylsis

Declining response rates have brought increased concerns from researchers, survey organizations, sponsors, and data users for the potential for nonresponse bias affecting survey estimates. Recently, the Office of Management and Budget (OMB) issued Standards and Guidelines for Statistical Surveys that encourages Federal agencies to plan for and conduct nonresponse bias analyses when their survey response rates fall below 80 percent. Thus, there is a growing need to identify and employ appropriate techniques to exam nonresponse bias across a variety of surveys. A panel of experts, Bob Groves (Univ. of Mich.), Mike Brick (Westat) and Lars Lyberg (Statistics Sweden), will share their expertise, insights, and experiments with different methods to estimate nonresponse bias. Brian Harris-Kojetin (OMB) will provide his insight and experience with developing and implementing the OMB guidance.

406 The 20/20 of Statistical Education: Where We've Been and Where We're Going

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Wednesday, August 1, 10:30 am–12:20 pm

The 20/20 of Statistical Education: Where We've Been and Where We're Going

❖Jeff Witmer, Oberlin College, Department of Mathematics, 205 King Building, Oberlin, OH 44074, *Jeff.Witmer@oberlin.edu*; ❖Allan Rossman, California Polytechnic State University, Department of Statistics, Faculty Office Building East (25), Room 102, San Luis Obispo, CA 93407, *arossman@calpoly.edu*; ❖Jessica Utts, University of California, Davis, Dept of Statistics, One Shields Ave, Davis, CA 95616, *jmutts@ucdavis.edu*; ❖George Cobb, Mount Holyoke College, Clapp Laboratory, Room 404C, South Hadley, MA 01075, *gcobb@mtholyoke.edu*

Key Words: Education, Challenges

We've made a lot of changes in statistical education in the last twenty years, but we still have a lot of work to do. Panelists will review changes that have been made and make projections for where we'll be in 5, 10, and even 20 years. They will also issue challenges to the statistical education community.

407 Seasonality and Statistics in Finance ●

Business and Economics Statistics Section

Wednesday, August 1, 10:30 am–12:20 pm

Seasonal Heteroscedasticity in Time Series: Modeling, Estimation, and Testing

❖William R. Bell, U.S. Census Bureau, SRD, Room 5K142A, 4600 Silver Hill Road, Washington, DC 20233-9100, *William.R.Bell@census.gov*; Thomas M. Trimbur, Federal Reserve Board

Key Words: seasonal ARIMA model, state space model, seasonal adjustment

Seasonal heteroscedasticity refers to regular changes in variability over the calendar year. Models for two different forms of seasonal heteroscedasticity were recently proposed by Proietti (2004) and Bell (2004). We examine use of likelihood ratio tests with the models to test for the presence of seasonal heteroscedasticity, and use of model comparison statistics (AIC) to compare the models and to search among alternative patterns of seasonal heteroscedasticity. We apply the models and tests to U.S. Census Bureau monthly time series of housing starts and building permits, and to time series of Industrial Production Indexes from the Federal Reserve Board.

A Local Goodness-of-Fit Diagnostic Based on the Log Determinant of the Sample Covariance Matrix

Tucker S. McElroy, U.S. Census Bureau; ❖Scott Holan, University of Missouri-Columbia, 3601 W Broadway Apt 30303, Columbia, MO 65203, *holans@missouri.edu*

Key Words: Frequency Domain, Nonstationary, Time Series

The log determinant of the covariance matrix of a stationary process has been used by various authors to provide information, and in particular can be used as a goodness-of-fit diagnostic on the residual sample autocovariances. This is equivalent to considering a total integral of the logged residual spectral density overall frequencies. We consider generalized goodness-of-fit measures that integrate over narrow bands of frequencies for use in assessing local goodness of fit. Additionally, we present theoretical and empirical results demonstrating the useful properties of the diagnostic. Finally, we illustrate the effectiveness of our diagnostic through application to real data from the U.S. Census Bureau.

Are Volatility Estimators Robust with Respect to Modeling Assumptions?

❖ Yingying Li, The University of Chicago, Department of Statistics, 5734 S University Ave, Chicago, IL 60637, *yyli@uchicago.edu*; Per A. Mykland, The University of Chicago

Key Words: Local Time, Market Microstructure, Martingale, Realized Volatility, Robustness, Two Scales Realized Volatility

We consider microstructure as an arbitrary contamination of the underlying latent securities price, through a Markov kernel Q. Special cases include additive error, rounding, and combinations thereof. Our main result is that, subject to smoothness conditions, the two scales realized volatility (TSRV) is robust to the form of contamination Q. To push the limits of our result, we show what happens for some models involving rounding (which is not, of course, smooth) and see in this situation how the robustness deteriorates with decreasing smoothness. Our conclusion is that under reasonable smoothness, one does not need to consider too closely how the microstructure is formed, while if severe non-smoothness is suspected, one needs to pay attention to the precise structure and also to what use the estimator of volatility will be put.

Covariance Estimation for Nonsynchronous High-Frequency Data with Microstructure Noise

❖ Qiuyan Xu, University of California, Davis, 421 Russell Park Apt 2, Davis, CA 95616, *qyxu@wald.ucdavis.edu*; Rituparna Sen, University of California, Davis

Key Words: Non-synchronicity, Microstructure noise, Random lead-lag estimator, Bias-variance trade-off

The use of high frequency return data has led to dramatic improvements in both theoretical and applied financial research. Covariance estimators among multiple processes have been proposed (e.g., De Jong and Nijman (1997) and Hayashi and Yoshida (2004, 2005, et al.). We introduce a new estimator, a random lead-lag estimator (RLLE) that coincides with the Hayashi-Yoshida estimator in the case where the frequency equals one second. We study the performance of RLLE for non-synchronous data under a situation with microstructure noise, and we obtain the optimal estimator with good bias-variance trade-off. Our result is confirmed by simulation results.

On Diagnostic Checking for Mixture Autoregressive Time Series

❖ Qin Shao, University of Toledo, Math Dept, Mailstop942, Toledo, OH 43606, *qin.shao@utoledo.edu*

Key Words: BIC, Empirical cumulative distribution function, Parametric bootstrap, Quantile-quantile plot

A diagnostic checking procedure is proposed for mixture autoregressive time series models. The asymptotic property is derived for the empirical cumulative distribution function of the residuals and is used to construct the quantile-quantile plot. The performance of this procedure is studied by simulation. The application is illustrated by two real time series datasets.

408 Interacting with the FDA in Medical Device Pre-Market Applications: From Pre-IDE to PMA Submission ● ✪

Biopharmaceutical Section, WNAR

Wednesday, August 1, 10:30 am–12:20 pm

Challenges in Trial Design for the Next Generation Drug-Eluting Stents

❖ Zheng Zhou, Boston Scientific Corporation, 100 Boston Scientific Way, M1 Biostatistics Dept, Malborough, MA 01752, *zhouf@bsci.com*; Peter Lam, Boston Scientific Corporation

Key Words: medical device, study design, regulatory submission

The presentation is focused on the drug eluting stent (DES) trial and its major hurdle after the first generation DES was approved in the US market. The strategies related to the trial design, the choice of study endpoint and the sample size for future generation DES trials are discussed. The interaction with the FDA statisticians in the approval process is crucial to the success of the IDE submission and a subsequent successful pre-marketing approval (PMA). Case examples are presented to share the experiences of interaction with FDA from the industry perspective.

Pre-PMA Meetings with the FDA in Medical Device Submissions

Greg Campbell, Food and Drug Administration; Gerry Gray, Food and Drug Administration; ❖ Lilly Yue, Food and Drug Administration, 1350 Piccard Dr., HFZ-550, Rockville, MD 20850, *lilly.yue@fda.hhs.gov*

Key Words: Premarket application, medical device statistics, diagnostic product

Prior to submitting a premarket application (PMA) to FDA's Center for Devices and Radiological Health (CDRH), a sponsor for a medical device or diagnostic product may choose to have a "pre-PMA" meeting with CDRH staff. This meeting can serve both regulatory and scientific purposes; CDRH statisticians regard it as an opportunity to discuss statistical issues with their industry counterparts, to set forth expectations regarding the submission of data in electronic form, and to arrive at understandings regarding data analysis and reporting. This talk will describe some of the outcomes of pre-PMA meetings that we feel are mutually beneficial to the sponsor's and FDA's statistical staff.

Some Thoughts on the Variables in Communications on Medical Device Studies Between the Sponsor and the FDA from the View of the Statistical Consultant

❖ Dennis King, STATKING Consulting, Inc., 759 Wessel Drive Unit 6, Fairfield, OH 45014, *statking@statkingconsulting.com*

Key Words: Medical Device, FDA, Statistical Consulting

During the period from Pre-IDE to PMA submission, the Sponsor and the FDA committees have many face-to-face and telephone communications concerning the Sponsor's medical device development program. Statisticians are employed by the Sponsor as the Sponsor's representative in the statistical matters related to the medical development program. The progress and effectiveness of these conversations is dependent on many variables including the level of involvement of the consultant in the Sponsor's company, the experience of the consultant and the experience level of the other

personnel (both Sponsor and FDA) involved in the communication. This talk will explore these and other variables in the communication process.

One Statistician's Thoughts on Nonstatistical Issues That Affect Statistical Considerations in Medical Device Studies

❖ Alicia Toledano, Biostatistics Consulting, LLC, 89 Hocken Avenue, Toronto, ON M6G 2K1 Canada, *toledano@rogers.com*

Key Words: advisory committee, indications for use, protocol agreement

Medical devices vary widely, and, partly because of this, so do recommendations made by FDA staff throughout the approval-seeking process. In the spirit of continuing dialogue with statistical consultants and FDA staff, I will share some perspectives from my experience. An underlying theme is maintaining professional integrity as one member of a larger group, in a situation that can truly make or break your client. Specific topics will include: how being a member of an FDA Advisory Committee helps in understanding the responsibilities of the statistical consultant; the importance of obtaining clear understanding and agreement between the company and the FDA on the proposed indications for use, and its impact on study design and end points; and the influence of who the FDA reviewers are, including their own experience, on protocol development and reaching a protocol agreement.

409 Topics in Crossover Trial Design and Analysis ●

Biopharmaceutical Section, WNAR, Section on Statistical Consulting, ENAR

Wednesday, August 1, 10:30 am–12:20 pm

Carryover Effects in Crossover Designs

❖ John Stufken, University of Georgia, Department of Statistics, Athens, GA 30602, *jstufken@uga.edu*

Key Words: change-over designs, repeated measurements

One of the most controversial topics when using a crossover design is how to deal with possible carryover effects. While there is broad agreement that one should attempt to use sufficiently long washout periods, this may not always be possible and it is not always clear what 'sufficiently long' really means. A growing number of researchers recommend that a crossover design should not be used if there is a possibility of carryover effects, with or without a washout period, that persist into the next active treatment period. Others have proposed more realistic models for modeling possible carryover effects. This talk will provide a closer look into these issues and review some of the proposals for alternative models.

Optimal and Efficient Crossover Designs for Comparing Test Treatments to a Control Treatment Under Various Models

❖ Min Yang, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65211, *yangmi@missouri.edu*; John Stufken, University of Georgia

We study the optimality, efficiency, and robustness of crossover designs for comparing several test treatments to a control treatment. Since A-optimality is a natural criterion in this context, we establish lower bounds for the trace of the inverse of the information matrix for the test treatments versus control comparisons under various models. These bounds are then used to obtain lower bounds for efficiencies of a design under these models. Two algorithms, both guided by these efficiencies and results from optimal design theory, are proposed for obtaining efficient designs under the various models.

Optimal Crossover Designs: Investigating the Role of Baseline Measurements

❖ Yuanyuan Liang, University of Alberta, CAB 632, Edmonton, AB T6G 2G1 Canada, *yliang@ualberta.ca*; Kimmie C. Carriere, University of Alberta

Key Words: optimal design, baseline measurements, self and mixed carryover effects model, random subject effects, Lagrange multiplier

In this paper, we apply the Lagrange multiplier method to solve the optimal design problem under a more general model: self and mixed carryover effects model with random subject effects. Generally, crossover designs lose the practical appeal when required to allow more than the traditional first-order carryover effects. This is when utilizing the baseline measurements clearly saves the design efficiency. We explore the effect of baseline measurements on optimal design results. We strongly recommend to use the baseline measurements as they improve the efficiency by 2-3 folds. We also find that there is a dramatic reduction in variability for estimating the direct treatment effect contrast when extending two-period designs to three-period or four-period designs. Such a reduction is much more pronounced when baseline measurements are incorporated in each period.

A Robust Bayesian Analysis of a Crossover Trial in Smoking Cessation

❖ Pulak Ghosh, Georgia State University, 30 Pryor Street, Atlanta, GA 30303, *pghosh@mathstat.gsu.edu*; Mary Putt, University of Pennsylvania

Key Words: bayesian, crossover design, Dirichlet Process prior, Gibbs sampling, smoking cessation

In a smoking cessation crossover trial, chronic smokers sequentially received placebo and naltrexone. Naltrexone is an opioid antagonist expected to reduce the reward value of nicotine, thus reducing motivation to smoke. In addition to an overall effect, differences in the responses of males and females to naltrexone were of interest. The distribution of two outcome variables reflecting positive and negative reinforcement from smoking displayed substantial evidence of departures from normality. Here we implement a Bayesian approach for estimating the effect of naltrexone on these outcomes. Our analysis is novel in the crossover literature in accounting for non-normality of the data through the use of a mixture of Dirchlet process priors on the model parameters. The approach is illustrated for each of the individual variables as well as in a multivariate framework.

410 Bayesian Applications Using Structured Stochastic Models ●✪

Section on Bayesian Statistical Science

Wednesday, August 1, 10:30 am–12:20 pm

Evolutionary Markov Chain Monte Carlo Algorithms for Expected Utility Maximization

❖ Marco Ferreira, University of Missouri-Columbia, Dept Statistics, 146 Middlebush Hall, Columbia, MO 65211-6100, *ferreiram@missouri.edu*; Ramiro Ruiz, Universidade Federal do Rio de Janeiro; Alexandra Schmidt, Universidade Federal do Rio de Janeiro

Key Words: Optimal Bayesian decisions, genetic algorithm, optimal design, environmental network design

We propose an evolutionary Markov chain Monte Carlo (EMCMC) framework for expected utility maximization. This is particularly useful when the optimal decision cannot be obtained analytically. Our proposed framework simultaneously computes the expected utility and maximizes it over the decision space. We develop an algorithm that simulates a population of Markov chains, each having its own temperature. The population evolves according to genetic operators, allowing the chains to explore the decision space both locally and globally. As a result, the algorithm explores the decision space very effectively. We illustrate that with two applications. First, we perform optimal design of a network of monitoring stations for spatio-temporal ground-level ozone. Second, we develop estimation of quantitative trait loci (QTL).

Bayesian Analysis of Inverse Gaussian Models for Response Time

❖ Paul Speckman, University of Missouri-Columbia, Department of Statistics, Columbia, MO 65211, *speckmanp@missouri.edu*; Yue Yu, University of Missouri-Columbia; Jeffrey Rouder, University of Missouri-Columbia

Key Words: Bayes, inverse Gaussian, psychology

Analysis of response time data is fundamental to research in cognitive psychology. In order to investigate cognitive theory, we propose response time models with parameters for shift, scale and shape. Determining the effect or lack of effect experimental manipulation has in each of these three parameters is crucial for theory development and validation. In this talk, we develop a hierarchical Bayesian analysis for a useful parameterization of the inverse Gaussian distribution with an added shift parameter. The method is illustrated with an application to a real data set, where we show that stimulus effects both scale and shape but not location. However, a re-parameterizing as a diffusion model shows that stimulus effects a single parameter, lending strong support to one particular cognitive model.

High-Resolution Space-Time Ozone Modeling for Assessing Trends

❖ Sujit Sahu, University of Southampton, School of Mathematics, Southampton, SO17 1BJ United Kingdom, *s.k.sahu@maths.soton.ac.uk*

Key Words: Dynamic model, misalignment, spatial variability, Markov chain Monte Carlo, stationarity

This paper proposes a space-time model for daily 8-hour maximum ozone levels to provide input for regulatory activities: detection, evaluation, and analysis of spatial patterns and temporal trend in ozone summaries. The model is applied to the analysis of data from the state of Ohio which contains a mix of urban, suburban, and rural ozone monitoring sites. The proposed space-time model is auto-regressive and incorporates the most important meteorological variables observed at a collection of ozone monitoring sites as well as at several weather stations where ozone levels have not been observed. This misalignment is handled through spatial modeling. In so doing we adopt a computationally convenient approach based on the successive daily increments in meteorological variables. The resulting hierarchical model is specified within a Bayesian framework and is fitted using MCMC techniques.

Bayesian Model Selection by Smoothly Clipped Continuous Priors

❖ Sounak Chakraborty, University of Missouri-Columbia, 134 K Middlebush Hall, Columbia, MO 65201, *chakrabortys@missouri.edu*

Key Words: model selection, SCAD, Bayesian analysis, classificartion, SVM

Model selection is present in the Bayesian literature for over a long time. The stochastic search variable selection or Variable selection through BIC has been extensively used for simultaneously choosing variables when we have a huge number of explanatory variables to decide from. A very apt example is the need for a sophisticated variable selection technique in gene expression microarray experiment. Under the regression setup the LASSO has been widely used for this purpose. Which has a direct relationship with a Bayesian model with double exponential prior. In this paper we talk about a new king of a continuous priors based on SCAD or smoothly clipped penalty function for variable selection. This SCAD prior can be used under both regression and classification setup and can be directly linked with Bayesian SVM for simultaneous classification and class prediction.

A Reference Prior for AR(2) Models

❖ Shawn Ni, University of Missouri-Columbia, Department of Economics, Columbia, MO 65211, *nix@missouri.edu*; Dongchu Sun, University of Missouri-Columbia

Key Words: reference prior, AR(2), MCMC

We develop a noninformative reference prior for AR(2) models that applies to stationary and nonstationary regions of the parameter space. We conduct MCMC simulations of the posterior under the reference prior. We also compare the finite sample properties of the posterior under the reference prior with those under several other noninformative priors, such as a flat prior on the AR(2)regression parameters and the Jeffreys prior.

411 Image Analysis in Solar and Astrophysics ● ✪

Section on Bayesian Statistical Science

Wednesday, August 1, 10:30 am–12:20 pm

Astronomical Imaging Using the GLAST Large-Area Telescope

❖ James Chiang, Stanford University, 2575 Sand Hill Rd, CRESST/UMBC, MS 29, Menlo Park, CA 94025, *jchiang@slac.stanford.edu*

The Large Area Telescope (LAT) aboard the GLAST satellite (to be launched Fall 2007) will survey the sky at very high photon energies, from 30 MeV to over 300 GeV. At these energies, conventional telescope mirror and detector technology cannot be used, and the directions and energies of incident photons are reconstructed by examining the properties of the particle showers produced in the LAT detectors. As a result, the blurring of the LAT is highly energy-dependent, with angular resolution going roughly as 1/E, and at low energies (< 200MeV), the point-spread function has an angular scale of a few degrees, with large non-Gaussian tails. This presents a severe source identification and image reconstruction challenge as we expect to observe many astronomical point sources clustered on scales smaller than this and extended sources that have structure on these angular scales and smaller.

Fully Bayesian Analysis of Low-Count Astronomical Images

❖ David van Dyk, University of California, Irvine, 2206 Bren Hall, Irvine, CA 92697-1250, *davd95@hotmail.com*; Alanna Connors, Eurika Scientific

Key Words: Astrostatistics, Multi-Scale Methods, Image Analysis, MCMC in Practice, Bayesian Methods, Poisson Models

New space-based telescopes that are designed to map X-ray and gamma-ray emission are giving a completely new perspective on the hot and turbulent regions of the universe. Analysis of the resulting images is a sophisticated task that requires subtle statistical techniques. Data is collected as photon counts on a grid of detector pixels. The counts are subject to non-uniform stochastic censoring, heteroscedastic errors in measurement, and background contamination. This combined with relatively small datasets makes answering complex astronomical questions a challenge. In this talk I describe how we (1) use Markov-random-field or multi-scale priors to stabilize the fitted images; (2) use posterior simulation to quantify uncertainty; (3) use higher resolution radio data to inform our priors; and (4) use Bayesian methods to test for deviations from particular structures in the image.

Multiscale Solar Image Processing

❖C. Alex Young, National Aeronautics and Space Administration, Goddard Space Flight Center, Building 26 Room 001 Mail Code 671.1, Greenbelt, MD 20771, *alex.young@gsfc.nasa.gov*; Jack Ireland, ADNET Systems, Inc.; R.T. James McAteer, Catholic University of America; Peter T. Gallagher, Trinity College Dublin

Key Words: wavelets, curvelets, image processing, astrophysics

Wavelets have been successfully used as a tool for noise reduction and general processing of images to view CMEs. Coronal Mass Ejections (CMEs) are large eruptions of plasma from the Sun. When CMEs travel to the Earth they can have hazardous effects including disrupting satellites and crippling electric power grids on the grounds. Because of this it is important to detect, track and understand CMEs to better predict their occurrence and impact on the Earth. The detection challenge is primarily because CMEs are extremely faint when compared to background. Wavelets have proved useful but have limitations. Wavelets are well suited for describing point singularities but edges or curves describe much of the interesting information in these images. The curvelet transform has promise as an optimal representation for CME images. We present a preliminary study of the use of these transforms.

Automatic Detection and Classification of Sunspot Images

❖Thomas C.M. Lee, The Chinese University of Hong Kong, Department of Statistics, The Chinese University of Hong Kong, Shatin, NT Hong Kong, *tlee@sta.cuhk.edu.hk*; C. Alex Young, National Aeronautics and Space Administration

Key Words: Astro-statistics, Morphology, Sunspots

In this work automatic methods are developed for detecting and classifying sunspots captured in images. An important element of our approach is the use of mathematical morphology. The good performance of our approach is demonstrated via real data examples. Future work include the tracking of detected sunspot features, which will play an important role in the understanding of the evolution of sunspot systems.

How to Find Loops in the Solar Corona

❖Vinay L. Kashyap, Harvard-Smithsonian Center for Astrophysics, 60 Garden St. MS83, Cambridge, MA 02138, *vkashyap@cfa.harvard.edu*; Julia Sandell, Barnard College; Thomas C. M. Lee, The Chinese University of Hong Kong

Key Words: AstroStatistics, Sun, X-ray, Morphological Analysis, Corona, Magnetic field

The Sun's corona is made up of hot plasma, typically at 1–2 million K. The corona is organized on large scales by the magnetic field of the Sun. Because the plasma is highly ionized, it forms loops tracing the field lines. The plasma cools by emitting high energy X-ray and EUV radiation, which is detected with the aid of orbiting solar telescopes. The detection and characterization of coronal loops thus leads to constraints on the magnetic and energetic environment on the Sun. However, coronal loops are notoriously difficult to extract from images because they overlap and form low-contrast images. Thus far, astronomers have depended upon visual inspection for the identification of interesting features. We have now implemented a morphological analysis process to identify and extract loop segments, and subsequently to model the coronal magnetic field using them. [NASA/LWSTRT:NNGO5GM44G]

412 Applications of Nonparametric Smoothing to Spatial and Temporal Data ● ✪

Section on Nonparametric Statistics, ENAR

Wednesday, August 1, 10:30 am–12:20 pm

Generalized Semiparametric Linear Mixed Effects Models

❖Tatiyana Apanasovich, Cornell University, 228 Rhodes Hall, School of ORIE, Ithaca, NY 14850, *tva2@cornell.edu*

Generalized linear mixed effects models are widely used for longitudinal non-Gaussian data analysis to incorporate between-subject and within-subject variations. To weaken model assumption for possible misspecification and to avoid the curse of dimensionality of fully nonparametric regression in the presence of several predictor variables, semiparametric models have been considered. We employ the penalized likelihood regression to estimate the models. We will focus on the efficient computation and the effective smoothing parameter selection. Real-data examples from AIDS studies will be presented to demonstrate the applications of the methodology.

Nonparametric Transfer Function Models: A Polynomial Spline Approach

❖Jun Liu, Georgia Southern University, 4325 Country Club Rd Apt 3, Statesboro, GA 30458, *jliu@georgiasouthern.edu*; Jing Wang, University of Illinois at Chicago

Key Words: nonparametric, transfer function, polynomial splines

Nonparametric transfer models have been developed to model nonlinear relationships between input and output time series. In this paper we consider a polynomial spline-based estimation method of such models. The transfer function is assumed to be smooth but the functional form unknown. The noise is assumed to follow a parametric ARIMA model. The transfer function is modeled using polynomial splines and estimated jointly with the ARIMA parameters. Compared with existing local polynomial-based approaches, the use of regression splines not only reduces the computational complexity, but also allows the noise to be nonstationary. The estimation procedures are introduced and the asymptotic properties of the estimators are discussed. The finite-sample properties of the estimators are studied through simulations and one real example.

Kernel Estimation of Multivariate Cumulative Distribution Function

❖Rong Liu, Michigan State University, Department of Statistics and Probability, East Lansing, MI 48824, *liurong@stt.msu.edu*; Lijian Yang, Michigan State University

Key Words: Berry-Esseen bound, GDP, kernel, rate of convergence, strongly mixing, survival function

A kernel estimator is proposed for multivariate cumulative distribution function, extending the work on Yamato (1973) on univariate distribution function estimation. Under assumptions of strict stationarity and geometrically strong mixing, we establish that the proposed estimator follows the same pointwise asymptotically normal distribution of the empirical cdf, while the new estimator is a smooth instead of a step function as the empirical cdf. Under stronger assumptions the smooth kernel estimator converges to the true cdf uniformly almost surely. Simulated examples are provided to illustrate the theoretical properties. Using the smooth estimator, survival curves for US GDP growth are estimated conditional on the unemployment growth rate to examine how GDP growth rate depends on the unemployment policy.

Lag Selection in Additive Autoregressive Models via Penalized Polynomial Spline

❖ Lan Xue, Oregon State University, Kidder 78, Corvallis, OR 97330, *xuel@stat.oregonstate.edu*

Key Words: Knot, Mean square consistency, penalized least squares, SCAD, US umemployment rate

A penalized polynomial spline estimation method is proposed for the nonlinear additive autoregressive models to simultaneously identify significant lags and estimate the corresponding additive nonparametric functions. The proposed method is a nonparametric extension of the nonconcave penalized least squares for linear models (Fan & Li 2001). It removes the redundant lags by shrinking small estimated nonparametric functions to zero, while estimates the functions of significant lags as accurate as the traditional polynomial spline method. Theoretical properties, such as, the existence and mean square consistency of the penalized polynomial spline estimators are established under strong mixing conditions. We demonstrate the good performance of the proposed method via comprehensive Monte Carlo studies and an application to the US unemployment rate.

413 The Statistician's Contribution to Fact-Finding in Law ●

Section on Statistical Consulting
Wednesday, August 1, 10:30 am–12:20 pm

Proof of Changepoints in Securities Prices Litigation: Use and Misuse of Event History Analysis

❖ Kenneth Swartz, Analysis and Inference, Inc., Victoria Mills, 1489 Baltimore Pike Suite 305, Springfield, PA 19064, *kswartz@analysisandinference.com*; Yulia R. Gel, University of Waterloo; R. Alan Miller, Philadelphia Investment Bank; William B. Fairley, Analysis and Inference, Inc.

Key Words: event history analysis, changepoint, time series, securities prices, lawsuit

The methodology of "event history analysis" treated in financial economics is often used for the purpose of testing the existence and estimating the amount of a changepoint in the price of a security. For example plaintiff shareholders bring a suit against a company whose stock appears to have suddenly dropped, and for which an explanation in terms of management actions is believed causal. Uses and misuses of this methodology, particularly from the viewpoint of time series analysis, are discussed with applications to cases.

Bagging and Propensity Scores in Statistical Learning to Assign Ownership of Unclaimed Property

❖ Marc Sobel, Temple University/Analysis and Inference, Inc., Victoria Mills, 1489 Baltimore Pike Suite 305, Springfield, PA 19064, *msobel@analysisandinference.com*; Kenneth Swartz, Analysis and Inference, Inc.

Key Words: bagging, propensity scores, cart, bootstrap, payment reversal

Statistical methods have been used to estimate the total amount putatively owed by a private entity to a government entity under unclaimed property law. A private entity may receive payments without identification amounting to millions of dollars. Since investigations are time-consuming, and in many instances are impossible to complete, a statistical method was sought that would provide a sound basis for classifying any unidentified payment as a mistaken double payment (lawfully becoming unclaimed government agency property) or not (a legitimate payment for services rendered by the corporation). We employ the procedure of bagging classification and regression trees (CART) using bootstrap resampling to identify such double payments. Bagging techniques have achieved a great deal of success in many applications.

Resolving a Multimillion Dollar Contract Dispute with a Latin Square Design

❖ Steven Crunk, San Jose State University, 375 Mullinix Way, San Jose, CA 95136, *crunk@math.sjsu.edu*; William B. Fairley, Analysis and Inference, Inc.; Marian Hofer, California State University, East Bay; Peter J. Kempthorne, Kempthorne Analytics, Inc.

Key Words: Latin Square, experimental design, contract, dispute, measure

A major eastern City negotiated a dispute over the performance of new garbage trucks with the vehicle manufacturer. The dispute concerned fulfillment of a contract calling for trucks to hold a "minimum full-load of 12.5 tons." On behalf of the City, and in cooperation with the manufacturer, statisticians developed a Latin Square design to estimate loads. The field experiment controlled for factors such as city districts, day-of-the-week, weight measurement method, instructions to crews, and supervision. To carry out the experiment, 5 pairs of trucks operated on 5 days in 5 districts. The contract spec did not further define the criterion of a "minimum full-load of 12.5 tons." Thus, different measures of contract fulfillment were defined. The experiment demonstrated successful fulfillment of the contract and resolved the dispute.

Statistical Assessment of the Fairness of a Firm's Allocation of Shares in Initial Public Offerings

❖ Efstathia Bura, George Washington University, Department of Statistics, 2140 Pennsylvania Avenue NW, Washington, DC 20252, *ebura@gwu.edu*

Key Words: IPO, fairness

In the late 1990s the stocks of technology companies, especially those involved with the Internet, were highly profitable. Customers of investment firms requested many more shares of new issues or initial public offerings (IPOs) than were available. In this talk, I will describe statistical measures for assessing the fairness of an investment firm's or broker's allocations of IPO shares to their customers. One approach compares the success rates of different groups of customers (e.g., those alleged to be favored with those of other groups). The second method incorporates the prior business each customer gave the firm into the comparison as an allocation criterion. This metric is an objective quantification of the fact that it is economically sensible for a firm to allocate more IPO shares to its 'best' customers.

414 Use of Measures Related to the Survey Variables in Post-Survey Adjustment for Nonresponse

Section on Survey Research Methods, Social Statistics Section
Wednesday, August 1, 10:30 am–12:20 pm

Preliminary Assessment of the Impact of Health Variables on Nonresponse Adjustment in the Medical Expenditure Panel Survey (MEPS)

❖ Trena Ezzati-Rice, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *trena.ezzati-rice@ahrq.gov*; Lap-Ming Wun, Agency for Healthcare Research and Quality

Key Words: nonresponse, propensity scores, weighting adjustments

MEPS is a complex sample survey sponsored by the Agency for Healthcare Research and Quality (AHRQ). Its sample is drawn from respondents to the previous year's National Health Interview Survey (NHIS). The MEPS, like most sample surveys, experiences unit nonresponse. Therefore, the base weight of MEPS respondents is adjusted to compensate for nonresponse. The first level adjustment is at the dwelling unit (DU) level to account for nonresponse among those households subsampled from NHIS for the MEPS. The adjustment is done using socio-economic, demographic, and health variables. Response propensity scores based on various combinations of covariates are calculated and quintiles of propensity scores are used to adjust the MEPS base weights for the assessment. This paper examines the impact on the DU level nonresponse weight adjustment with and without inclusion of health variables.

Using Interviewer Observations To Improve Nonresponse Adjustments: NES 2004

❖ Andy Peytchev, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709-2194, *apeytchev@rti.org*; Kristen Olson, University of Michigan

Key Words: nonresponse, weighting adjustment, propensity scores

One widely used method for adjusting for the effects of unit nonresponse is weighting. Propensity models and weighting class adjustments commonly employ available demographic characteristics. However, interviewers can make observations about the sample member, the housing unit, and the doorstep interaction that could be related not just to the propensity to respond, but also to statistics of interest. Using data from the 2004 American National Election Study we demonstrate whether the use of such interviewer-collected auxiliary data is useful for weighting adjustment beyond the current poststratification weights. In addition we demonstrate how adjustments that use interviewer observations affect only some weighted estimates. We conclude with a discussion of the need to collect auxiliary variables that are associated with key statistics of interest, not only response propensity.

Use of Interviewer Judgments About Attributes of Selected Respondents in Post-Survey Adjustment for Unit Nonresponse: An Illustration with the National Survey of Family Growth

Robert M. Groves, University of Michigan; James Wagner, University of Michigan; ❖ Emilia Peytcheva, University of Michigan, 2722 Packard Street, Apartment H, Ann Arbor, MI 48108, *emilia@umich.edu*

Key Words: nonresponse, postsurvey adjustment, propensity models, paradata

With computer-assisted data collection methods, the survey designer has the ability to seek interviewer observations about sample persons. When such observations can be made on both respondents and nonrespondents, the resulting paradata are candidates for use in adjustment. This paper evaluates the measurement error properties of an interviewer observation about attributes related to key measures in the National Survey of Family Growth (NSFG), sponsored by the US National Center for Health Statistics. Immediately after the household roster is taken, one age-eligible respondent is selected using fixed probabilities based on age and gender. Prior to any initiation of the main interview request, the interviewer is asked to judge attributes of the sample person. These are evaluated using a variety of post-survey adjustment approaches and a variety of key statistics computed on the NSFG.

Using Proxy Measures of Survey Outcomes in Post-Survey Adjustments: Examples from the Consumer Expenditure Survey (CES) and the Los Angeles Neighborhood Observations Study (LA FANS)

❖ Frauke Kreuter, University of Maryland, 1218 Lefrak Hall, JPSM, College Park, MD 20742, *fkreuter@survey.umd.edu*; Carolina Casas-Cordero, University of Maryland

Key Words: weighting adjustments, nonresponse adjustments, response propensity, consumer expenditure survey

With the goal of removing bias due to nonresponse, survey researchers often create post-survey nonresponse adjustment weights. To be effective, variables used need to be highly correlated with the response propensity and the survey variable of interest. Correlates of key survey variables can often be collected at almost no additional cost for researchers. One example are interviewer observations during the field period. The Bureau of Labor Statistics already includes interviewer observation of housing tenure status in the CES adjustment procedure. Using data from the 2004 CES we estimate propensity models with and without the interviewer housing tenure variable and compare weights created out of those propensity models. We will develop a similar analysis using the LA FANS study, which provides a unique set of physical and social observations recorded by interviewers at the block level.

Using Proxy Measures of the Survey Variables in Post-Survey Adjustments in a Transportation Survey

❖ Ting Yan, University of Michigan, Institute for Social Research, 2019 Medford Road, Apt J270, Ann Arbor, MI 48104, *tyanuconn@gmail.com*; Trivellore E. Raghunathan, University of Michigan

Key Words: post-survey adjustments, nonresponse error, proxy measures of survey variables, response propensity

Nonresponse bias and the associated sampling error are reduced when the post-survey adjusted estimate employs variables that are related to both the survey variable of interest and propensity to respond. Using the survey conducted by the University of Michigan Transportation Research Institute, this paper will take advantage of the driving history data obtained from the State of Michigan and develop a new weighting scheme drawing on proxy measures of the survey variables, design variables, and process measures that are highly correlated with both response propensity and the survey variable of interest. This paper will then compare this new weighting scheme against a more traditional one employing only demographic variables. Both weighting methods will be assessed on their effectiveness in reducing nonresponse bias.

415 Computationally Intensive Methods in AIDS Research I ● ✪

WNAR, Biometrics Section, ENAR, Section on Health Policy Statistics, Section on Statistics in Epidemiology

Wednesday, August 1, 10:30 am–12:20 pm

Mathematical Modeling and Statistical Methods for AIDS Translational Research from Benchside to Bedside

❖ Hulin Wu, University of Rochester, Dept of Biostatistics and Computational, 601 Elmwood Ave Box 630, Rochester, NY 14642, *hwu@bst.rochester.edu*

Key Words: HIV/AIDS, Viral Dynamics, Differential Equation Models, Viral Fitness

We propose statistical methods to estimate HIV viral dynamic parameters in both in vitro experiments and clinical trials that are described by a set of nonlinear ordinary differential equations (ODE). It is quite challenging to study the identifiability and fit the nonlinear differential equation models without a closed form solution. We employ an engineering technique to study the identifiability of the proposed ODE models. For model fitting and parameter estimation, we propose the naive nonlinear least squares method that requires numerically solving the ODEs and a two-step method that avoids the numerical solution of the ODEs. We present the asymptotic results of the proposed estimators and finite-sample properties via Monte Carlo simulations. We also illustrate the proposed methods by applying the methods to HIV viral fitness experimental data and AIDS clinical trial data.

A Multiscale Model of HIV Dynamics Linking In-Host Processes with Virus Spread in Human Populations

❖ Lilit Yeghiazarian, University of California, Los Angeles, Department of Biostatistics, CHS 51 239 D, Los Angeles, CA 90025, *lyeghiaz@ucla.edu*; William G. Cumberland, University of California, Los Angeles; Otto Yang, University of California, Los Angeles

Key Words: multi-scale, sexual networks, model, in-host process

Interactions of HIV with the immune system have been studied experimentally and using mathematical models. Epidemiological studies provided insight into HIV dynamics in human populations. The importance of social network topology in epidemics became apparent as research on structure of technological networks progressed. We have developed a mathematical model that bridges the scales between in-host processes and HIV transmission in human sexual networks. Each individual goes through a sequence of health states reflecting HIV status and treatment. Associated with each state is an in-host model of HIV-immune system interactions. At the same time, each individual is a node in a sexual network. We investigate the effect of a variety of medical care criteria and decisions on both the health state of the individual, and on the spread of disease in populations with different network topologies.

A Stochastic Model of HIV Escape from the Cytotoxic T Lymphocyte Response

❖ Elissa Schwartz, Washington State University, Department of Mathematics, PO Box 643113, Pullman, WA 991643113, *schwartz@math.wsu.edu*; Lisette G. De Pillis, Harvey Mudd College; William G. Cumberland, University of California, Los Angeles; Otto Yang, University of California, Los Angeles

Key Words: Stochastic Model, Parameter Estimation, Human Immunodeficiency Virus, Vaccine, Cytotoxic T Lymphocyte Response, Viral Escape

HIV vaccine development requires information on what stimulates an effective immune response. Determination of conditions allowing viral evasion from killing by cytotoxic T lymphocytes (CTLs) is needed in order to predict patterns of viral escape. The appearance of CTL escape mutants reported in the published data is not well approximated by deterministic models of viral escape. This finding motivated us to model viral escape as a stochastic process (taking into account viral parameters such as mutation rate, production rate, and the effect of CTL killing), fit the models with clinical data, and estimate kinetic parameters involved in escape. Our calculations may be used to provide estimates of the CTL pressure on HIV. Furthermore our results underscore fundamental immune mechanisms that may help design prophylactic and therapeutic HIV vaccines.

Application of a Model with Two Types of CTL Regulation to Shiv Vaccination Trials

❖ Igor Rouzine, Tufts University, Microbiology, 136 Harrison Ave, Boston, MA 02111, *irouzine@tufts.edu*

Key Words: vaccine, helper, cytotoxic, model

Recently, we proposed a mathematical model of interaction between HIV and immune system that explained a score of quantitative observations made in HIV/SIV infection. The model included cytotoxic lymphocytes (CTLs) activated directly by virus, and those activated by virus-specific CD4 helper cells. Some viruses (HIV/SIV, LCMV) can infect helper cells. The model predicts the existence of two steady states: one with a high virus load, few helper cells, and moderate CTL number (usually observed); and another with a very low virus load, and high numbers of helper cells and CTLs. Precisely tuned external intervention, such as antiretroviral therapy or vaccine, in principle, can switch the high-virus state to the low-virus state. Here, we predict the vaccine-induced number of memory CTLs required to reach the low-viremia state for animals vaccinated against the recombinant virus SHIV.

Two-Way Hierarchical Phylogenetic Models: An Application to HIV-1 Evolution During Fusion Inhibitor Therapy

❖ Christina Ramirez Kitchen, University of California, Los Angeles, 650 Charles E Young Dr, Dept. Biostatistics, Los Angeles, CA 90095-1772, *cr@ucla.edu*; Marc A. Suchard, University of California, Los Angeles

Key Words: Bayesian, phylogenetic, heirarchical, HIV, T20, genetics

We studied intra-host viral evolution during T20 treatment during advanced HIV disease. We examined clones from 9 patients with incomplete viral suppression on T20, and who interrupted T20 while remaining on background antiretrovirals. Clones were sequenced from gp41 and gp120 at three time points: pre-T20, post-T20 failure and post-T20 interruption. To test the hypothesis of dominant selective pressure simultaneously in gp41 and gp120 across patients, we extended the Bayesian hierarchical phylogenetic model to a two-way ANOVA design. The first dimension is across patients and the second is across gene regions. The hierarchical prior on patient trees can be thought of as a two-way contingency table with each cell having multinomial probabilities under a log linear formulation. This model can also be used for Bayesian outlier detection.

416 Gene Expression and Genomics ● ✪

Biometrics Section

Wednesday, August 1, 10:30 am–12:20 pm

Genome Phylogeny Accommodating Lateral Gene Transfer

❖ Hongmei Zhang, University of West Florida, Dept Math and Stat, 11000 University Parkway, Pensacola, FL 32514, *hzhang@uwf.edu*; Rohan Hemasinha, University of West Florida; Xun Gu, Iowa State University

Key Words: Lateral gene transfer, Extended gene content, Maximum likelihood, Stochastic process, Genome distance

This paper developed an approach to inferring gene evolution parameters and general genome distances under gene loss, birth, and lateral gene transfer using extended gene content. Simulation results indicated that, even with small number of gene families, with extended gene content information general genome distances under LGT can be well estimated using the developed. We applied the new method to 13 genome microbial complete genomes. By comparing the results from the new approach to those from others, the new method is useful to study the universal tree of life and also to reveal the evolution pattern.

A Simulation Study of the Operating Characteristics of Statistical Analysis Plans for Genome-Wide Association Studies

❖John Barnard, The Cleveland Clinic, Quantitative Health Sciences Wb4, 9500 Euclid Ave, Cleveland, OH 44195, *barnarj@ccf.org*

Key Words: multi-locus, missing data, power, multiple comparisons

Genome-wide association (GWA) studies are a potentially powerful approach for detecting genetic associations with complex disease phenotypes. How to best analyze a GWA study remains an open and active area of research. We have conducted numerous simulation studies that attempt to estimate the frequentist operating characteristics of various statistical analysis plans. These plans cover (1) choice of stages; (2) quality control practices; (2) missing data treatment; (4) population stratification approaches; (5) multiple comparison procedures; and (6) choice of test statistics. Simulated data were generated from artificial populations constructed by in-silico random mating starting with a founder population consisting of real GWA data. We will present a summary of the simulation studies and conclusions about recommended statistical analysis plans for GWA studies.

Analyzing Data from Gene Expression in Neuronal Activity

❖Xiaofeng Wang, The Cleveland Clinic, Department of Quantitative Health Scie, 9500 Euclid Ave, Cleveland, OH 44195, *wangx6@ccf.org*

Key Words: Spatial count data, Over-dispersion, Simultaneous inference, Neuronal gene expression, Neuroscience

Neuronal gene expression is known to play a key role in determining the differentiation of neuronal development. An understanding of how neuronal activity modifies gene expression can help elucidate mechanisms of long-term neuronal plasticities in the brain. Data from gene expression in neuronal activity are commonly spatially-correlated counts of activated neurons. We discuss several models for the complex inhomogeneous count data: from parametrics to semiparametrics. Simultaneous confidence regions in these models are investigated. A real example in neuroscience are also explored in the paper.

Mutation Detection in Exon Grouping Analysis

❖Yuan Yuan, Harvard University, 1 Oxford St, 7th floor, Cambridge, MA 02138, *yyuan@stat.harvard.edu*; Xin Lu, University of California, San Diego; Jun S. Liu, Harvard University; James D. Iglehart, Dana-Farber Cancer Institute; Alexander Miron, Dana-Farber Cancer Institute

Key Words: DNA mutation, Exon Grouping Analysis, outlier detection

The analysis of DNA for mutations based on resulting structural aberrations has been performed in various forms for decades. DNA can be analyzed after it has been denatured to a single strand form by a procedure called Single-Strand Conformational Analysis (SSCA). DNA that contains a mutation folds in a conformation that differs from that of wild type DNA. These fragments can be analyzed on a polyacrylamide gel and mutations can be identified due to their aberrant mobility. We have developed a process that is high throughput, very sensitive as compared to standard Sanger sequencing (98% sensitivity in a blinded validation study) and inexpensive. A nonparametric approach combining nearest-neighbor and quantile statistic is applied to select outlier DNA fragments which may contain mutation. The result may also be applied to discover Single Nucleotide Polymorphisms (SNPs).

Gene-Gene and Gene-Environment Interactions for Complex Traits

❖Qinyi Cheng, New York University, 650 1st Ave Rm 581, New York, NY 10016, *qc218@nyu.edu*; Mengling Liu, New York University; Yongzhao Shao, New York University

Key Words: gene-gene interaction, gene-environment interaction, association study, Ancestry informative marker

In studies of gene-gene and gene-environment interactions for complex traits, latent population stratification usually entails complications. Unobserved differences in subpopulation proportions among cases and controls may lead to false positive results. To incorporate potential latent population substructure into the analysis, we propose a joint model of phenotypic traits and genotypic information under the presence of genetic information of Ancestry-Informative-Markers (AIMs). Moreover, numerical algorithms to carry out the estimation of relevant parameters are developed. Simulation studies are used to demonstrate the effect of taking consideration of population admixture.

Characterization of Transcriptomes by Application of Reliability Analysis in Structural Genomics

❖Fisseha Abebe, Clark Atlanta University, 223 James Brawley Dr, Atlanta, GA 30314, *fabebe@cau.edu*; William Seffens, Clark Atlanta University

Key Words: mrna, randomized sequence, weibull random variable, folding energy

Bioinformatic studies have demonstrated that local mRNA stability is important in gene expression. Previous examination of 51 mRNAs revealed a bias toward more negative folding energies in gene sequences globally. Instead of examining a limited number of genes, this statistical study considered about 6000 Human, 12000 Mouse and 17000 Arabidopsis complete sets of mRNAs termed transcriptomes and analyzed the variation of energy values among genes of differing length. Randomized sequence folding free energies were found to have the same distribution but with means having different slopes for the three transcriptomes. Further investigation showed that the Weibull random variable describes mRNA fold distribution within and among genomes more accurately than conventional power functions which have been used in a number of structural genomic studies.

417 Applications of Semiparametric and Nonparametric Modeling Techniques ● ✪

Biometrics Section, Section on Nonparametric Statistics

Wednesday, August 1, 10:30 am–12:20 pm

Weighted Fourier Series Analysis of Anatomical Brain Structures

❖Shubing Wang, University of Wisconsin-Madison, Department of Statistics, 1245 Medical Sciences Center, 1300 University Ave, Madison, WI 53705, *shubing@stat.wisc.edu*

Key Words: Weighted, AIR, registration

As the generalized forms of Fourier Series and Spherical Harmonics, WFS are superior since they not only effectively reduce Gibb's phenomenon, but provide a convenient way to compute smoothness when applying random field theories. An Adaptive Iterative Regression (AIR) method is proposed for WFS representations of large-sized data, which preserves the computational efficiency of previous iterative fitting methods and improve the computational efficiency of Least-Squared Estimations. Using a novel method of calculating curvatures, we propose a curvature-based global shift registration method. Decision-tree-based classification techniques and random field theories to study the link between autism and its underlying neuroanatomy.

A Semiparametric Empirical Likelihood Method for ROC Curve Regression of Ordinal Diagnostic Test in the Presence of Verification Bias

❖ Xiaofei Wang, Duke University, Medical Center, Room 8027, Hock Plaza, 2424 Erwin Road, Durham, NC 27717, *xiaofei.wang@duke.edu*; Haibo Zhou, The University of North Carolina at Chapel Hill

Key Words: Empirical Likelihood, Ordinal Diagnostic Test, Receiver Operating Characteristic (ROC) Curve, Semiparametric Method, Verification bias

In medical studies to assess the accuracy of diagnostic tests, definitive disease assessment is too invasive or expensive to be ascertained on all the study subjects. It is often more ethical or cost effective to ascertain the true disease status with a higher rate in study subjects where the diagnostic test is suggestive of disease. Without taking the study design into account, standard statistical methods can yield biased estimates for diagnostic accuracy. It is known as verification bias in the literature. To accommodate the biased sampling scheme, we develop a semiparametric empirical likelihood method to estimate the parameters of the covariate-specific ROC regression model for ordinal diagnostic test. We consider biased sampling schemes in which the set of selected subjects include either a simple random sample, or a test-result-dependent sample or both.

A Novel Wavelet Approach for Denoising Mass Spectrometry Data

❖ Jaesik Jeong, Texas A&M University, 1100 Hensel Dr Apt U1K, College Station, TX 77840, *jjeong@stat.tamu.edu*

Key Words: Maximal overlap discrete wavelet transform, mass spectrometry, heteroscedastic error model, SELDI TOF M/S

In recent years, there has been an increased interest in using protein mass spectroscopy to detect discriminating molecular markers. One of popular analyses in proteomics is based on mass spectrometry data. When we analyze this kind of data, it is necessary to preprocess data. With poorly preprocessed data we have minuscule chance to find biomarkers. Wavelet denoising technique becomes one of standard approach to preprocessing procedure. However, existing wavelet denoising technique does not reflect the existence of heterogeneity error in mass spectrometry data. In this paper, we propose a novel wavelet denoising to deal with the heterogenity.

Nonparametric Quantification of Protein Lysate Arrays

❖ Jianhua Hu, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 237, Houston, TX 77030, *jhu@mdanderson.org*

Key Words: Cobs, logistic modeling, nonparametric quantification, protein lysate array

Proteins play a crucial role in biological activity, so much can be learned from measuring protein expression directly. The Reverse-phase protein lysate arrays allow us to quantify the relative expression levels of a protein in many different cellular samples simultaneously. Existing approaches to quantify protein arrays use parametric response curves fit to dilution series data. The results can be biased when the parametric function does not fit the data. It is difficult to model all the samples and proteins with a convenient parametric response curve. We propose a nonparametric approach which adapts to any monotone response curve. The nonparametric approach is shown to be promising via both simulation and real data studies. The nonparametric approach enables more reliable quantification of protein lysate arrays.

Semiparametric Two-Sample Changepoint Model

❖ Zonghui Hu, National Institutes of Health, 6700A Rockeldge Drive, Room 5231, Bethesda, MD 20817, *huzo@niaid.nih.gov*; Jing Qin, National Institutes of Health; Dean Follmann, National Institute of Allergy and Infectious Diseases

Key Words: changepoint, empirical maximum likelihood, likelihood ratio, semiparametric

A semiparametric changepoint model is proposed to investigate the possible difference between two distributions. The empirical maximum likelihood (EML) method is adopted for the estimation of the changepoint and the two marginal distributions. It is shown that the EML estimator of the change-point is consistent. The likelihood ratio test for the existence of changepoint is provided, as well as the goodness-of-fit test on the changepoint model. We provide an application in a clinical study to evaluate the equivalence of two devices for measuring the viral load. Though the two devices were generally considered as equivalent for underlying viral load of large values, we showed that the equivalence fails for low viral loads.

Double Nonparametric Estimation for Relating Dose Distributions to Scalar Outcomes

❖ Matthew Schipper, University of Michigan, 200 Zina Pitcher Place, 3307A Kresge III, Ann Arbor, MI 48109, *mjschipp@med.umich.edu*; Jeremy Taylor, University of Michigan

Key Words: functional data, monotonicity, nonparametric regression, dose effect, normal tissue complications

Normal tissue complications are a common side effect of radiation therapy. They are the consequence of the non-uniform dose of radiation received by the normal tissue near the tumor. Dose-damage-injury models relate the dose distribution received by the normal tissue to the observed injury/complication in two steps. The first step relates dose to the unobserved damage and the second relates damage to the observed injury. In our model, a summary measure of damage is obtained by integrating a weighting function of dose (W(d)) over the dose density. Similar to a generalized additive model, the linear predictor in our model includes a nonparametric function of damage, H(damage) which relates damage to injury. Both W(.) and H(.) are written as regression splines and estimated nonparametrically and monotonically. We illustrate our method with data from a head and neck cancer study.

Asymptotic Theories in Multiple-Sample Semiparametric Density Ratio Models

❖ Guanhua Lu, University of Maryland, Math Dept., 8261 Canning Ter, Greenbelt, MD 20770, *ghlu@math.umd.edu*; Benjamin Kedem, University of Maryland

Key Words: Semiparametric, Density Ratio Model, Biased Sampling, Weak convergence of stochastic processes, Asymptotic theory, Gaussian process

A multiple-sample semiparametric density ratio model can be constructed by multiplicative exponential distortions of the reference distribution. Distortion functions are assumed to be nonnegative and of a known finite-dimensional parametric form, and the reference distribution is left nonparametric. The combined data from all the samples are used in the large sample problem of estimating each distortion and the reference distribution. The large sample behavior for both the parameters and the unknown reference distribution are studied. The estimated reference distribution is proved to converge weakly to a zero-mean Gaussian process. And the corresponding covariance structure is used to provide the confidence bands. A Kolmogorov-Smirnov type statistic is also studied for a goodness-of-fit test of the density ratio model.

418 Adaptive Group Sequential Designs in Drug and Vaccine Development ● ✪

Biopharmaceutical Section

Wednesday, August 1, 10:30 am–12:20 pm

A Hybridized Two-Stage Design in an Ulcerative Colitis Pivotal Trial

❖ Guohui Liu, Millennium Pharmaceuticals, Inc., 40 Landsdowne Street, Cambridge, MA 02139, *guohui.liu@mpi.com*; Mingxiu Hu, Millennium Pharmaceuticals, Inc.

Key Words: two-stage design, type I error, bias, power

Ulcerative colitis (UC) is an inflammatory disease. A novel humanized antibody is under development as a treatment for active UC patients. This paper presents a hybridized two-stage design for a pivotal study to prove the efficacy of the antibody in the induction of remission of UC and to establish the optimal dose. Stage 1 serves as the confirmatory study for the overall population combining anti-TNF naÔve and refractory populations, and also as the learning stage for anti-TNF refractory population. Stage 2 proceeds as the confirmatory study for the anti-TNF refractory population only if Stage 1 demonstrates a treatment effect for the subpopulation that passes a pre-specified threshold. Sample size for Stage 2 is based on Stage 1 results on the subpopulation. A simulation study has been done to explore the characteristics of this design including overall type I error, bias, and power.

Multiplicity Adjustment for Testing of Multiple Endpoints in Group Sequential Study Design

❖Jianming Wang, Amgen Inc.; Xiang Zhang, Amgen Inc., 2811 Arguello Drive, Burlingame, CA 94010, *xiangz@amgen.com*; Chunlei Ke, Amgen Inc.

Key Words: hierarchical testing, Group sequential study design, primary endpoint, secondary endpoints, type I error rate

Group sequential study design with multiple endpoints is very common in clinical trials, and the group sequential method is widely used in the primary efficacy endpoint in terms of controlling type I error rate. Not many papers have been published regarding how to control type I error rate using a group sequential rule based on multiple endpoints. In particular, when primary and secondary endpoints are tested hierarchically (secondary endpoints will be tested only if primary endpoint is tested significantly), it is not quite clear what is the optimal in controlling type I error rate in testing secondary endpoints. The purpose of our research is to evaluate performances of several common procedures in terms of controlling type I error rate strongly (experimental-wise error rate), and wish to shed some lights on how to make selection among several common procedures.

Self-Designing Multiple Dose Study via Adaptive Randomization

❖ Lin Wang, sanofi-aventis, 200 Crossing Blvd BX2 409A, Bridgewater, NJ 08807, *lin.wang2@sanofi-aventis.com*; Lu Cui, Eisai Medical Research Inc.

Key Words: clinical trials, adaptive randomization, type I error rate

In clinical trials, in addition to the control, multiple doses of a testing new drug are often used for various purposes, including the determination of a dose response curve or of a potentially more effective dose for further investigation. In this study, a new adaptive randomization method is proposed to improve the efficiency of a multiple dose trial. The new method allows unequal patient allocations, as driven by the data, to different treatment arms according to the trial objectives. The descriptions of the new method will be given together with the assessment of its performance in terms of the power, the type I error rate, and the rate of successful patient allocations. Its implementation, requiring no Bayesian calculations, as well as its applications to Phase II dose response studies and Phase III studies with dropping ineffective dose arms will also be discussed.

Estimation Following an Adaptive Group Sequential Test

❖ Cyrus Mehta, Cytel Inc., 675 Massachusetts Avenue, Cambridge, MA 02139, *mehta@cytel.com*

Key Words: adaptive clinical trial, parameter estimation, adaptive inference, confidence intervals, phase 3 adaptive trial

This paper proposes two methods for computing confidence intervals with exact or conservative coverage following a group sequential test in which an adaptive design change is made one or more times over the course of the trial. The key idea, due to Muller and Schafer (2001), is that by preserving the null conditional rejection probability of the remainder of the trial at the time of each adaptive change, the overall type 1 error, taken unconditionally over all possible design modifications, is also preserved. This idea is further extended by considering the dual tests of repeated confidence intervals (Jennison and Turnbull, 1989) and of stage-wise adjusted confidence intervals (Tsiatis, Rosner and Mehta, 1984). The method extends to the computation of median unbiased point estimates.

A Supremum Log-Rank Test for Adaptive Two-Stage Treatment Strategies and Corresponding Sample Size Formula

❖ Wentao Feng, University of Pittsburgh, 3754 Allendale Circle, Pittsburgh, PA 15204, *wef5@pitt.edu*; Abdus S. Wahed, University of Pittsburgh

Key Words: Adaptive treatment strategies, Counting process, Proportional hazards, Sample size formula, Supremum log rank statistics, Survival Function

In two-stage adaptive treatment strategies, patients receive one of the induction treatments, followed by a maintenance therapy given that the patient responded to the induction therapy. To test the difference in the effect of different induction and maintenance treatment combinations, a modified supremum weighted log-rank test is proposed. The test is applied to the data from a two-stage randomized trial for illustration and compared with standard weighted log rank test. A sample size formula is proposed based on the limiting distribution of the supremum weighted log-rank statistic. Monte Carlo simulation studies show that the proposed test provides sample sizes which are close to that obtained by standard weighted log-rank test under proportional hazard alternative. The test is more powerful than standard weighted log-rank test under non-proportional hazard alternatives.

Sample Size Re-estimation for Adaptive Sequential Design in Clinical Trials

❖ Ping Gao, The Medicines Company, 1019 Chambers Ct, Bridgewater, NJ 08807, *ping.gao@themedco.com*; James H. Ware, Harvard University

Key Words: sequential design, sample size re-estimation, conditional power, conditional type I error, Brownian motion

In designing clinical trials, sample size is often estimated based on prior information of the effect of the test drug or test treatment regimen, in the form of expected size of the treatment effect and its variance. It is therefore often desirable to perform an interim analysis or interim analyses to reassess the treatment effect to either stop the trial early due to overwhelming evidence of efficacy or increase sample size to assure adequate power. We propose a conditional power based sample size calculation for sequential designs. Our method extends Fisher's (1998) result, which does not allow for early termination.

On Group Sequential Adaptations in Vaccine Development

❖William Wang, Merck & Co., Inc., 351 North Sumneytown Pike, Mail stop: UG1CD-38, North Wales, PA 19454, *william_wang@merck.com*; Ivan Chan, Merck & Co., Inc.; Joseph Heyse, Merck & Co., Inc.

Key Words: Vaccine Trials, Group Sequential Adaptation

The application of adaptive designs has recently been studied and intensely debated in the development of drug products. In this talk, we will point out that group sequential adaptations could readily be applied to vaccine development due to many unique features. These features include (1) relatively short risk period in vaccine safety evaluation (2) convenient immune marker in vaccine immunogenicity trials, and (3) long-term follow-up in vaccine efficacy trials. We further discuss sample size re-estimation, alpha spending strategy, adjusted group sequential estimation, and two-stage mid-course adaptation in the context of vaccine development. We will use two recently completed pivotal vaccine trials, the Shingles Prevention Trial (Oxman et al., NEJM 2005) and the Rotavirus Efficacy and Safety Trial (REST) (Vesikari et al, NEJM 2006), along with simulations to illustrate our points.

419 General Statistical Issues and the Role of Statisticians in Medical Research ● ✪

Biopharmaceutical Section, Biometrics Section

Wednesday, August 1, 10:30 am–12:20 pm

The Role of the Statistician in Biomedical Research: Scientist, Shoe Clerk, or Hired Gun

❖ Anthony Segreti, RTI Health Solutions, 3040 Cornwallis Road, Research Triangle Park, NC 27709, *asegreti@rti.org*

Key Words: Ethical Guidelines, Role of Statistician, Biomedical Research

JAMA recently stated that it would no longer accept industry-sponsored trials in which the data analysis was performed solely by the sponsor's statisticians, instead requiring that the analysis be performed by or independently validated by an academic statistician. Although this policy has created some concern in the statistical community, the reaction has been modest. I suggest that this policy has important implications for the role of the statistician in biomedical research and deserves more attention that it has received so far. Irwin Bross wrote a 1974 article on the role of the statistician. He suggested that statisticians in biomedical research had to decide if they were willing to take on the full responsibility of a scientist. This question continues to be relevant today but the statistician must also deal with the scrutiny of critics like JAMA.

Comparison of Predictive Values of Two Diagnostic Tests Using Conditional Model

❖Yoonjin Cho, North Carolina State University, 1408 crest Rd 103, Raleigh, NC 27606, *einsheart@gmail.com*; Andrzej S. Kosinski, Duke University

Key Words: Positive predictive value, Negative predictive value, Diagnostic test, Conditional model, Logistic regression

Positive and negative predictive values are important measures when one compares performance of two diagnostic tests. Thus, the development of statistical methodology to test equality of these measures between two diagnostic tests is essential. In this presentation we consider a paired design in which each subject has both tests performed. We first review the existing methods for comparison of predictive values in such a design. The first is based on marginal regression and the second is weighted least squares. We propose a conditional model approach. Our model uses logistic regressions and Wald statistic is used to test equality of positive or negative predictive values between the two diagnostic tests. We conduct simulations under a variety of scenarios to evaluate type I error and power of the proposed testing approach and compare to the other two existing methods.

Statistical Performance of Cladistic Strategies for Haplotype Grouping in Pharmacogenetics

❖Jared Lunceford, Merck & Co., Inc., 4087 N Foothill Dr, Provo, UT 84604, *jared_lunceford@merck.com*; Nancy Liu, Merck & Co., Inc.

Key Words: pharmacogenetics, haplotype grouping, cladistic analysis, randomized trials

Haplotypes comprised of multiple single nucleotide polymorphisms (SNPs) are popular covariates for capturing the key genetic variation present over a region of interest in the DNA sequence. The multiallelic nature of haplotypes increases the complexity of the statistical models intended to discover association with outcomes of interest. Cladistic methods for haplotype grouping cluster haplotypes according to estimates of their genealogical closeness and have been proposed recently as approaches for reducing model complexity and increasing power. Two such approaches are compared in the context of a case study: methods based on a haplotype nesting algorithm described by Templeton (1987) and hierarchical clustering of haplotypes as described by Durrant et al. (2004).

Statistics in Translational Medicine

❖Dennis Cosmatos, Wyeth, 500 Arcola Rd, Rm A3154, Collegeville, PA 19426, *cosmatd@wyeth.com*; Shein-Chung Chow, Duke University

Key Words: Translational Medicine, Translational Science, One-way Translation, Two-way Translation, Predictive modeling, Lost in translation

Pizzo(2006) defines translational medicine as bench-to-bedside research where in a basic laboratory discovery becomes applicable to the diagnosis, treatment or prevention of specific diseases and is brought forth by either a physician-scientist who works at the interface between the research laboratory and patient care or by a team of basic and clinical science investigators. Statistics plays an important role in translational medicine to ensure that the translational process is accurate and reliable with certain statistical assurance. For this purpose, statistical criteria for assessment of one-way and two-way translation are proposed. Under a well established and validated translational model, statistical tests for one-way and two-way translation are derived. Statistical inference for loss in translation and for the applicability of an animal model to a human model are also discussed.

Working for a Medical Device Company versus a Pharmaceutical Firm

❖Yanping Chang, Medtronic Inc., 1015 Gramsie Road, Mail Stop Z220, Shoreview, MN 55311, *yanping.chang@medtronic.com*

Key Words: Medical Device Company, Parmaceutical Firm

Statisticians can apply their knowledge of statistics to a variety of fields such as public health, medicine, economics, engineering, education, and sports, and work in government, business and industry. Many statisticians work in medical device and pharmaceutical companies to develop and evaluate the results of clinical trials to determine the safety and effectiveness of new medications or medical devices. However, there are differences working in medical device companies and pharmaceutical firms. 1.The differences of devices and drugs. 2. Statistical organizational difference. 3. Statistician role and responsibilities (i.e., data cleaning, database management, SAS programming, and report writing). 4.Clinical trial design and applied statistical knowledge (i.e., Bayesian statistics in medical device trials). 5.Outsourcing.

Assessment of the Placebo Effect in Quality-of-Life Evaluations

❖Jens Eickhoff, University of Wisconsin-Madison, 610 Walnut St 250 WARF, Madison, WI 53726, *eickhoff@biostat.wisc.edu*

Key Words: Cardiovascular clinical trial, mixture modeling, EM-algorithm

The assessment of Quality of Life (QoL) has become a standard component in many clinical trials. QoL, however, is an inherently subjective measure and as such, very sensitive to the placebo effect. The placebo effect has led to dramatic consequences on the clinical development of new therapeutic agents. A mixture regression model is proposed to assess the placebo effect in a placebo-controlled clinical trial. Simulation studies are conducted to examine the empirical behavior of the proposed method. The approach is illustrated with the analysis of QoL data from a phase III cardiovascular clinical trial.

On Errors in Stratifications

❖Chunlei Ke, Amgen Inc., One Amgen Center Drive, MS 24 2 C, Newbury Park, CA 91320, *cke@amgen.com*; Jianming Wang, Amgen Inc.; Xiang Zhang, Amgen Inc.; Qi Jiang, Amgen Inc.; Alan Rong, Amgen Inc.; Steven Snapinn, Amgen Inc.

Key Words: stratified randomization, clinical trial, error in variable, chi-square test

Stratified Randomization is widely used in clinical trials to achieve balance of the treatment assignment with regard to important prognostic factors. It may happen that the value on the stratification factor provided at randomization is later identified to be incorrect late. Thus patients may be allocated into the strata they should not belong to. We evaluated the impact of such errors on the intended balance, and found that randomization was still balanced at least at the level as in simple randomization. While the analysis based on the factor per randomized is usually chosen, the estimate of treatment effect based on such analysis was found unbiased on average. But the inference based on this based on this analysis approach was shown conservative, and thus inefficient. In contrast, the analysis utilizing the true stratification factor was found proper and more efficient.

420 Methods for Survival Analysis ● ✪

ENAR, Section on Risk Analysis

Wednesday, August 1, 10:30 am–12:20 pm

Application of Survival Models to Vaccine Efficacy Evaluation

❖Barbara Krasnicka, Food and Drug Administration, 1401 Rockville Pike, CBER, Rockville, MD 20852, *barbara.krasnicka@fda.hhs.gov*; Sang Ahnn, Food and Drug Administration

Key Words: vaccine efficacy, survival analysis, Cox model, predictive accuracy, heterogeneity

Essential problems encountered in vaccine clinical trials are: finding to what extent the vaccine prevents the targeted disease, and establishing a correlation between disease protection and antibody responses as measured at a given time point after vaccination. In the past, to address these issues, models based on survival analysis approaches, in particular Cox proportional hazard regression model, were utilized. However, interpretation of Cox model parameters is limited. Therefore, there is a need to apply methods such as evaluations of explained variation and/or predictive accuracy. In our presentation, we focus on vaccine efficacy estimation based on time-to-event data, in particular on handling heterogeneity problems in such data. Some useful methods will be discussed.

An Adjusted Cumulative Incidence Function Estimator for Unbalanced Groups

❖Yongyun Zhao, University of Pittsburgh, Department of Biostatistics, 311 Parran Hall GSPH, 130 Desoto Street, Pittsburgh, PA 15261, *yoz6@pitt.edu*; Chung-Chou H. Chang, University of Pittsburgh

Key Words: Adjusted cumulative incidence function, competing risks, cumulative incidence function, inverse probability weighting,

Cumulative incidence function (CIF) estimates are commonly used to describe the failure probability when competing risks are present. CIF estimates can be derived from Kaplan-Meier survival and cause-specific hazard estimates. In observational studies or nonrandomized trials, the CIF estimates may be biased if the distribution of a confounding variable differs among groups. To reduce bias, we developed an adjusted CIF (ACIF) estimator that is based on the use of inverse probability weighting. To obtain the ACIF, we first determine the probability that each individual will be in a particular group, conditional on the individual's characteristics. After we assign a weight equal to the inverse of this probability, we use the weight to adjust the Kaplan-Meier survival and cause-specific hazard estimates. To illustrate, we use the example of HIV treatment groups.

Robust Estimation in the Presence of Extreme Censoring

❖Ruta Bajorunaite, Marquette University, PO Box 1881, Milwaukee, WI 53201, *ruta@mscs.mu.edu*; Vytaras Brazauskas, University of Wisconsin-Milwaukee

Key Words: survival data, censoring, parametric models, robust estimation

Parametric models are frequently used in modeling survival data. In many studies we encounter data where a certain proportion of extreme observations is censored. Maximum likelihood estimation is a standard tool used to fit parametric models. However, when the underlying model is mis-specified or contaminated the maximum likelihood parametric methods may be severely affected and lead to very poor results. We propose a robust parametric model fitting procedure for continuous failure time models. The procedure is based on trimmed L-statistics and it can achieve various degrees of robustness, which can be easily specified by the user as trimming proportions. Unlike most M-estimators, the newly developed method is straightforward to implement in practice as it (typically) does not require numerical solution of non-linear equations. The procedure is illustrated using real data example.

Methods for Cox Regression with Nonclassical Measurement Error in the Covariates with an Application to Nutritional Epidemiology

❖Pamela A. Shaw, National Institutes of Health, 6700 A Rockeldge Drive, Room 5230, Bethesda, MD 20817, *shawpa@niaid.nih.gov*; Ross L. Prentice, Fred Hutchinson Cancer Research Center

Key Words: Measurement Error, Survival Analysis, Nutritional Epidemiology, Cox Regression

Self-reported measures of diet are error prone, yet most historical analyses relating diet to disease have not adjusted for systematic aspects of measurement error. We develop methods to analyze censored failure time data with Cox regression when the covariate of interest has both subject-specific and systematic measurement error. Three techniques for the classical error model, riskset regression calibration (Xie, Wang, and Prentice, JRSSB, 2001), conditional score (Tsiatis and Davidian, Biometrika, 2001), and nonparametric corrected score (Huang and Wang, JASA, 2000), are extended to the case where the covariate of interest follows a generalized measurement error model and a covariate measured with unbiased error is available on a subset. An example from the Women's Health Initiative motivates this more flexible model and provides the basis of a simulation study to compare the methods.

Modeling Associations via Incidence Functions with Bivariate Competing Risks Data

❖Yu Cheng, University of Pittsburgh, 2717 Cathedral of Learning, Department of Statistics, Pittsburgh, PA 15260, *yucheng@pitt.edu*; Jason Fine, University of Wisconsin-Madison

Key Words: Cause-specific failure patterns, Empirical processes, Nonparametric estimation, Time-dependent association

Frailty models are frequently used to analyze clustered survival data and evaluate within-cluster associations. However, they are seldom used in bivariate competing risks settings because of the multiple interacting failure types. To address this issue, we focus on a nonparametrically identifiable quantity: cumulative incidence function (CIF). Frailty models are constructed expressing the bivariate CIF in terms of its marginals based on some improper random variables. Estimating equations are proposed to estimate the unknown association parameter involved in frailty models. The large sample properties of the association parameter estimators are established using empirical processes techniques and their practical performances are studied by Monte-Carlo simulations. We illustrate their practical utility by an analysis of dementia in the Cache County Study.

A Generalized Bivariate Gamma Distribution and Its Characterization

❖Norou Diawara, Old Dominion University, 4700 Elkhorn Ave, 2317 ECS Building, Norfolk, VA 23529, *ndiawara@odu.edu*; Mark Carpenter, Auburn University

Key Words: Survival, Reliability models, Bivariate Gamma, Dirac measure

Because of the numerous applications, characterization of multivariate survival distributions is still of a growing research area. We restrict the marginals to a specified lifetime distribution, while proposing a linear relationship between the marginals with an error random variable that we completely characterize. The distributions are all of positive support, but some one class has a positive a probability of simultaneous occurrence. In that sense, we capture the absolutely continuous case, and the Marshall-Olkin type with a positive probability of simultaneous event on a set of measure zero. We extend the results available in the literature and propose and study the properties of various parameter estimators.

A Taxonomy of Hazards for the Generalized Gamma Distribution

❖Christopher Cox, Johns Hopkins Bloomberg School of Public Health, 615 N Wolfe St Room E7642, Dept. of Epidemiology, Baltimore, MD 21205, *ccox@jhsph.edu*; Haitao Chu, Johns Hopkins Bloomberg School of Public Health; Michael Schneider, Johns Hopkins Bloomberg School of Public Health; Alvaro Munoz, Johns Hopkins Bloomberg School of Public Health

Key Words: survival analysis, parametric models, generalized gamma distribution, non-proportional hazards, relative times

The generalized gamma (GG) offers flexibility for analysis of censored and truncated survival data. We present a GG hazard taxonomy. The half plane of the scale and shape parameters is divided into four regions within which hazards are bathtub and arc shaped, increasing and decreasing. The GG includes the Weibull, lognormal and gamma, and is closed under power transformations so inverse distributions are included. Characterization of the quantile function allows use of relative quantiles as effect measures. We apply the taxonomy to describe survival after clinical AIDS during different eras of HIV therapy, with hazard functions of very different shapes. Standard errors for relative times and hazards are computed using the delta method and checked with the bootstrap. These methods allow cohort studies to provide assessment of treatments at the population level of public health relevance.

421 Bayesian Applications in Genomics ● ✪

Section on Bayesian Statistical Science

Wednesday, August 1, 10:30 am–12:20 pm

Bayesian Pathway Annotation Analysis of Genomewide Expression Profiles

❖Haige Shen, Duke University, 214A Old Chemistry Bldg, Computational Biology and Bioinformatics, Durham, NC 27708-0251, *haige.shen@duke.edu*; Mike West, Duke University

Key Words: gene set enrichment analysis, marginal likelihood, pathway annotation, variational approximations

Understanding molecular pathways underlying cancer phenotypes is essential to uncovering dynamic process of tumorigenesis. As part of this, linking quantified, experimentally defined gene expression signatures with known biological pathway gene sets is a key challenge. We develop a Bayesian approach to the problem of genome-wide expression-based pathway annotation. This involves a model-based approach to matching experimental signatures of structure or outcomes in gene expression—represented in terms of ranked and weighted gene lists—to multiple pathway gene sets from curated databases. One overall result is posterior probabilities over pathways for each experimental signature. We discuss the modeling approach and some computational challenges, and demonstrate the use of MCMC and variational methods that provide solution. Examples in cancer pathway analysis highlight the approach.

A Bayesian Approach to Pool Multiple Datasets for Periodically Expressed Genes Detection

❖ Xiaodan Fan, Harvard University, 1 Oxford Street, 7th FL, Cambridge, MA 02138, *xfan@fas.harvard.edu*; Jun S. Liu, Harvard University

Key Words: reversible jump MCMC, Bayesian meta-analysis, cell cycle gene, transformation group, time series, model selection

Multiple microarray time series datasets have been produced in an effort to detect periodically expressed genes. There are discrepancies with regard to the identity and numbers of periodically expressed genes. We proposed a hierarchical model to pool all datasets to get better classification of genes as either periodically expressed (PE) or aperiodically expressed (APE). A damping sinusoidal function with linear tread and iid Gaussian noise is used to model each PE time series. APE time series are simple modeled by a linear tread plus iid Gaussian noise. Within an experiment, all genes share the same period. Between experiments, all genes share the same phase shift. MCMC is used to estimate the parameters. Reversible Jump MCMC is used to dynamically select models (PE or APE) for each time course. Transformation move, group move, block swapping are used to improve the mixing of the chain.

Bayesian Nonparametric Clustering of Amino Acid Usage Profiles

❖ Daniel Merl, Cornell University, 580 Malott Hall, Ithaca, NY 14853, *dmm229@cornell.edu*

Key Words: Dirichlet process, positive selection, Bayesian nonparametrics

The study of heterogeneity in amino acid usage is central to the field of molecular evolution. Increased heterogeneity at individual amino acid sites can be the result of relaxation of constraint, or the result of recurrent adaptive evolution (positive selection). I describe a Dirichlet process mixture model for achieving Bayesian nonparametric clustering of amino acid sites according to their usage profiles. Identification of sites with high posterior probabilities of belonging to clusters representing increased heterogeneity levels is shown by simulation study to be useful for detecting possible targets of positive selection. Since the clustering induced by the Dirichlet process is based on an unknown number of mixture components, the method is appropriate for the analysis of protein coding data for which there is little prior information about the variation in selective pressure.

A Fast Bayesian Method for eQTL Linkage Analysis in Experimental Crosses

❖ Jinze Liu, The University of North Carolina at Chapel Hill, 132 Ridge Trail, Biostatistics Department, Chapel Hill, NC 27516, *liuj@cs.unc.edu*; Fred Wright, The University of North Carolina at Chapel Hill; Fei Zou, The University of North Carolina at Chapel Hill; Yu-Ling Chang, The University of North Carolina at Chapel Hill

Key Words: eQTL, Baysian, Linkage analysis, Laplace approximation

We apply a recently-developed approximate Bayesian linkage analysis approach to the expression quantitative trait loci (eQTL) problem, in which microarray measurements of thousands of transcripts are examined for linkage to genomic regions. The approach uses the Laplace approximation to integrate over genetic model parameters (not including genomic position), and has been fully developed for different types of recombinant inbred crosses. The method is much faster than commonly-used Monte Carlo approaches, and thus suitable for the extreme computational demands of eQTL analysis. We have formulated biologically attractive priors involving explicit hyperparameters for probabilities of cis-acting and trans-acting QTLs. The approach offers highly interpretable direct posterior densities for linkage for each transcript at each genomic position.

Bayesian Variogram Modeling in a Metric Space

❖ Yan Zheng, sanofi-aventis, 1212 Cornerstone Blvd, Apt 327, Downingtown, PA 19335, *yan.zheng@sanofi-aventis.com*; Cavan Reilly, The University of Minnesota

Key Words: Bayesian inference, gene expression, Bessel functions, correlation functions, variograms, Gibbs' sampler

We propose an approach to Bayesian variogram modeling for data that is referenced to a metric space and apply the method to the analysis of Affymetrix oligonucleotide arrays. Before fitting a parametric model to the correlation structure, we investigated the possibility of anisotropy. We mapped the oligonucleotide space to a metric space using multidimensional scaling method and the directional semivariograms are used to assess anisotropy. We investigate the impact of the sequence on the measured intensity and fit a 25-way ANOVA to the intensity with factors being the nucleotide in each of the 25 positions of the 25mer. Finally we computed the residuals from the model and fit Bessel mixture Bayesian variogram models for different number of Bessel mixture components on these residuals.

A Bayesian Model for Genetic Association with Phenotype Defined in a Limited Range

❖ Ling Wang, Boston University, 401 Broadway, Apt 35, Cambridge, MA 02139, *wangling@bu.edu*; Vikki Nolan, Boston University; Clinton Baldwin, Boston University; Martin Steinberg, Boston University; Paola Sebastiani, Boston University

Key Words: Bayesian, DIC, Beta distribution, SNP

The phenotype from genetic association studies is often a continuous variable such as a probability, or a percentage, and are defined in [0,1]. Linear regression with the transformed phenotype as dependent variable may lead to false positive associations and produce poor fit due to the transformation. We propose a Bayesian approach that uses Beta distribution to model the phenotype in the correct range of definition. We estimate the association with or without adjustment for confounding in a Bayesian framework using MCMC computation, and evaluate this procedure by assessing the false positive rates (FPR) and true positive rates (TPR) in simulated data. Compared with the model assuming a lognormal distribution, our method has a comparable FPR but a higher TPR. We apply this method to candidate-gene analysis that might impact fetal hemoglobin concentration in sickle cell anemia patients.

Two-Way Clustering of DNA Microarray Data Using the Bayesian Plaid Model

❖ Jiajun Gu, Harvard University, 36 Oxford Street, Room 416, Cambridge, MA 02138, *jiajungu@fas.harvard.edu*; Jun Gu, Harvard University

Key Words: Bayesian plaid model, bi-clustering, normalization of DNA microarray data, gibbs sampling

In recent years, two-way clustering or bi-clustering algorithms are in particular interest for DNA microarray analysis. A bi-cluster refers to a subset of genes and a corresponding subset of experimental conditions of a microarray dataset. In this work, a Bayesian bi-clustering model is proposed. Similar to the plaid model of Lazzeroni and Owen in 2002, the Bayesian plaid model uses an ANOVA model to represent a bi-cluster and allows an unknown number of bi-clusters. Gibbs sampling algorithm is used to search for multiple bi-clusters simultaneously. We applied the model to both simulated and real datasets including leukemia and yeast microarray data and identified bi-clusters with significant biological meanings. We also compared it with a few published bi-clustering algorithms based on their operation characteristics such as specificities, sensitivities and overlapping rates.

422 Design and Analysis of Experiments ● ✪

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Wednesday, August 1, 10:30 am–12:20 pm

Nested Space-Filling Designs for Experiments with Multiple Levels of Accuracy

❖Zhiguang Qian, University of Wisconsin-Madison, Department of Statistics, 1300 University Ave, Madison, WI 53706, *zhiguang@stat.wisc.edu*

Key Words: space-filling designs, design of experiments, computer experiments, discrete mathematics, finite fields, orthogonal arrays

Computer experiments with different levels of accuracy have become prevalent in many engineering and scientific applications. Design construction for such experiments is a new issue because traditional methods deal almost exclusively with experiments with one level of accuracy. In this work, we construct some nested space-filling designs for experiments with different levels of accuracy. Our construction makes use of Galois fields and orthogonal arrays.

Mixed Effects Models for Prediction of Tool Wear

❖ Arzu Onar, St. Jude Children's Research Hospital, 332 N Lauderdale St, Biostatistics Mailstop 768, Memphis, TN 38105, *arzu.onar@stjude.org*

Key Words: unobserved heterogeneity, factorial experiment, longitudinal data, degradation

Semi-dry and dry machining is being adopted by the metal cutting industry worldwide to reduce the harmful effects of traditional metal cutting fluids and the cost associated with procurement, use, and disposal of these fluids. A factorial experiment was designed to study the effect of speed, feed, depth of cut and cutting condition on tool wear in a machining context where tool wear progression was measured longitudinally. The resulting data was modeled via mixed effects to identify the main effects as well as the interactions which play a part in tool wear progression.

Geometric Equivalence of Symmetric Factorial Designs

❖ Tena Katsaounis, The Ohio State University, 1185 Yorkwood Rd, Mansfield, OH 44907, *katsaounis.1@osu.edu*; Angela Dean, The Ohio State University; Cheryl Dingus, Battelle

Key Words: geometric equivalence, symmetric factorial design, experimental design

In this talk, the problem of geometric equivalence of symmetric factorial experiments is considered. Two fractional factorial symmetric designs with quantitative factors are said to be geometrically equivalent if one can be obtained from the other by reordering the runs, relabeling the factors and reversing the symbols of factors levels. Existing criteria for detecting geometric equivalence or nonequivalence of symmetric factorial designs are described and evaluated via computer algorithms. Some new necessary and sufficient criteria of equivalence are presented. All results generalize to designs with factors having different number of levels.

Comparison of Slack Variable and Mixture Approaches for Mixture Experiments

❖ Greg Piepel, Battelle/PNNL, PO Box 999, Richland, WA 99352, *greg.piepel@pnl.gov*

Key Words: Mixture experiment, Slack variable, Component, Experimental design, Model

A mixture experiment involves blending two or more components in various proportions and measuring one or more response variables on each blend. The slack variable (SV) approach is one possible approach for designing and modeling the data from mixture experiments. With the SV approach, one component is designated the SV, the experiment is designed, and the data are analyzed in terms of the remaining components. A SV proportion is obtained by subtracting from one the sum of the proportions of the remaining components, thus "taking up the slack." With the mixture approach, the experiment is designed and the data are analyzed using all of the mixture components. The presentation considers the situations in which the SV approach has been proposed, and argues in each situation that it is generally preferable to use an appropriate mixture approach.

Optimal Foldovers of Minimum Aberration Even Designs

❖ Robert W. Mee, University of Tennessee, 333 Stokely Management Center, Department of SOMS, Knoxville, TN 37996-0532, *rmee@utk.edu*; Jihua Xiao, University of Tennessee

Key Words: design augmentation, fractional factorial, resolution IV, semifolding

Montgomery and Runger (1996) discuss the best foldover plans for resolution IV fractional factorial designs for up to 10 factors. Here we present general results for identifying the best foldover for resolution IV designs for k factors in n = 2k runs and projections of these designs. We also consider semifolding, since follow-up designs with n/2 runs exist that permit estimation of as many effects as are provided by a regular foldover.

A Class of Optimal Three-Level Response Surface Designs

❖ William Li, The University of Minnesota, 321 19th Ave S, Operations and Management Sci Dept, Minneapolis, MN 55311, *setosa@yahoo.com*; Xuan Lu, Tsinghua University

Key Words: augmented pairs design, central composite design, optimal design, orthogonal design, response surface design, sequential experiment

Three-level experimental designs are popular choices for fitting response surface models. Morris (2000) proposed a class of three-level composite designs, which are constructed sequentially. At stage 1, a two-level orthogonal design is selected; and at Stage 2, augmented points resulting from pairs of points of the two-level designs are added. The resulting designs are called augmented pairs (AP) designs. We consider optimal AP designs by finding the best two-level orthogonal such that the resulting three-level AP designs have optimal criteria values. Both orthogonal and non-orthogonal two-level designs are considered. In the second part of the research we construct minimal AP designs, for which the two-level designs used in Stage 1 are restricted to be saturated designs.

Projection Properties of Hadamard Designs and Their Application

❖ Yingfu Li, University of Houston-Clear Lake, 2700 Bay Area Blvd, 182, Houston, TX 77058, *li@uhcl.edu*

Key Words: Hadamard design, projection, robust parameter design

In a screening design, often only a few factors from a large number of potential factors are significantly important. The projection property is an important issue when selecting a screening design. In this talk, we shall discuss the projection properties of Hadamard designs with run-size up to 60. We are typically interested in finding Hadamard designs whose all $k-$column projections are able to entertain and estimate all 2-factor-interactions along with its main effects, assuming higher order interactions are negligible. A discussion is also made of the possible application of our finding to the robust parameter design, an important statistical tool in quality engineering with aims at reducing the variation of the system.

423 Estimation

IMS, Section on Statistical Computing
Wednesday, August 1, 10:30 am–12:20 pm

A Generalized P-Value Approach to Inference on Common Mean

❖ Pranab K. Mitra, Bristol-Myers Squibb Company, 912 C Hooper Ave, Baltimore, MD 21229, *pranab1@umbc.edu*

Key Words: Common Mean, Generalized p-value, Coverage Probabilities

Statistical inference for the common mean of several univariate normal populations has many useful applications in practice. In this paper, we develop a generalized p-value approach to derive tests and confidence intervals for the common mean based on several well known estimates of the common mean. The case of a multivariate normal distribution with a common component mean is also addressed and a generalized p-value approach based on the maximum likelihood estimate of the common mean is developed. We also report simulated powers and coverage probabilities based on the proposed methods.

Implementation of an Extended Familial Correlation Structure with Quasi-Least Squares

❖ Jichun Xie, University of Pennsylvania, School of Medicine, 5th Floor Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, *jichun@mail.med.upenn.edu*; Justine Shults, University of Pennsylvania

Key Words: quasi-least squares, qlsinr, extended familial data

The method of quasi-least squares (QLS, Chaganty and Shults, 1999) is an approach in the framework of GEE that allows for easier implementation of the extended familial correlation structure. We discuss this structure and its implementation in an analysis of a study in Ophthalmology using the qlsinr function for R software. An important goal of our presentation is to explain and demonstrate implementation of qlsinr, so that audience members will be able to implement this free software in their own analyses.

On the Asymptotic Properties of the Signed Composite Likelihood Ratio Statistic

❖ Zi Jin, University of Toronto, 100 St George St, Department of Statistics, Toronto, ON M5S 3G3 Canada, *jin@utstat.toronto.edu*; Nancy Reid, University of Toronto

Key Words: Composite likelihood, Pairwise likelihood, Asymptotic efficiency, Higher order asymptotics, Signed composite likelihood ratio statistic

Composite likelihood has received increased attention in handling large dimensional data sets when the computation of the full likelihood is highly complex. In particular, we concentrate on pairwise likelihood, which arises by considering two-dimensional marginal densities. The pairwise likelihood estimators are generally consistent and asymptotically normally distributed. We summarize different test statistics, as well as their asymptotic distributions. Higher order asymptotic properties of the signed composite likelihood ratio statistic are outlined. Some examples are analyzed in detail, and simulation studies are presented to further explore the asymptotic performance of the composite likelihood ratio statistic.

Preferred Distribution Function for Estimating P(Y< X)

❖ Mohammed A. Shayib, Prairie View A&M University, Department of Mathematics, P O Box 519, Prairie View, TX 77446-0519, *mashayib@pvamu.edu*; Aliakbar Haghighi, Prairie View A&M University

Key Words: Reliability, Probability Distributions, MLE, Populations' Parameters, Samples' sizes, Simulation

The estimation of $P(Y < X)$, assuming X and Y to be independently and identically distributed random variables, has been extensively studied in the literature. Distribution functions such as Normal, Burr Type X, Exponential, Gamma and Weibull have been considered. In this paper, we consider same type of distributions. Utilizing the Maximum Likelihood Estimation method and using simulation we will estimate the unknown parameters of these distributions. To choose a preferred distribution, comparison among results will be made based on the ratio of the parameters and the sample size.

A Stationary Stochastic Approximation Algorithm for Estimation in Generalized Linear Mixed Models

❖ Sheng-Mao Chang, North Carolina State University, 2300 Avent Ferry Road, Apt D4, Raleigh, NC 27606, *schang2@unity.ncsu.edu*; John F. Monahan, North Carolina State University

Key Words: stochastic approximation, RCA, mixed model, top Lyapunov exponent

Estimation in generalized linear models is challenging because the likelihood is an integral without closed form. Among those leading solutions the likelihood is approximated and the maximum likelihood estimate (MLE) can only be reached with error. The simultaneous perturbation stochastic approximation (SPSA) algorithm is designed to find the exact MLE under the same circumstances but provides no error bound if the algorithm is stopped in a finite steps. In order to estimate MLE properly with an error bound, we design the stationary SPSA (SSPSA) algorithm. Assuming that the marginal likelihood is quadratic around the MLE, the SSPSA takes the form of a random coefficient vector autoregressive model. Under some mild conditions, the algorithm yields a stationary sequence where the mean of this sequence is asymptotically unbiased and consistent to the MLE.

Generalized Estimating Equations for Nonlinear Mixed Effect Model

❖ Lawrence Lee, University of California, Davis, Department of Statistics, One Shields Avenue, Davis, CA 95616, *llee@wald.ucdavis.edu*; Jiming Jiang, University of California, Davis

Key Words: Gaussian quadrature, Laplacian approximation, Linear mixed-effects approximation, generalized estimation equation

Nonlinear mixed effect models (NMEM) are widely used in pharmacokinetics. Currently, the most popular approaches in inference about NMEM are: Linear mixed-effects approximation, Laplacian approximation and Gaussian quadrature. While these approaches are quite effective computationally, they are known to produce inconsistent estimators of the parameters. We propose a generalized estimation equation approach to the parameter estimation in nonlinear mixed effect model and study its performance by Monte Carlo simulations.

424 Surveillance, Identity, and Other Issues in Defense Statistics ●

Section on Statisticians in Defense and National Security

Wednesday, August 1, 10:30 am–12:20 pm

Military Requirements Forecasting: A Reliability-Based Approach

❖ Dennis Mathaisel, Babson College, Babson Hall 317, Babson Park, MA 02457-0310, *mathaisel@babson.edu*

Key Words: Military Requirements Forecasting, Reliability-Based Forecasting

The US military has been experiencing severe problems due to inaccuracies in their requirements forecasting: substantial budget increases from unanticipated demands; parts and materials shortages; and low safety stock levels. The Requirements Definition Forecasting process is critical to military sustainability. Improving RDF accuracies will greatly enhance operations by minimizing material and parts shortages and help improve warfighter operations by correctly predicting maintenance and logistics requirements necessary in meeting specified readiness levels. Current methods use traditional statistical demand forecasting methodologies based upon historical trends. This paper investigates a reliability-based approach that integrates empirical field data into physics of failure models to predict end-of-life failure modes and rates.

An Optimization Approach for Identity Matching in the Federal Sector

❖ David Lloyd, MITRE Corporation, 7515 Colshire Dr, McLean, VA 22102-7508, *dalloyd1@cox.net*; Keith Miller, MITRE Corporation

Key Words: identity, matching, entity resolution, search algorithms, search performance metrics, name matching

The approach incorporates the use of ground-truthed control data sets along with the use of sample sets of adjudicated matching tests carried out by separate matching validation groups as a basis for the computation of separate F scores for each search algorithm used. An overall F score is then calculated from the application of all combinations of search algorithms against a target set of queries used with the target identity database comparable to the database used in the Federal Agency. Although the approach is modified to include the use of other identity attributes in order to optimize large-scale processing, the primary intent of the approach is to reduce the risks associated with the name-matching component of the identity matching process.

Addressing the Problem of Data Scarceness in Risk Analysis with Hierarchical Coordinated Bayesian Models

❖ Zhenyu Yan, University of Virginia, 547 Seymour Road Apt 3, Charlottesville, VA 22903, *zyan@virginia.edu*; Yacov Y. Haimes, University of Virginia

Key Words: Risk analysis, Hierarchical Bayesian Model, Markov chain Monte Carlo Simulation

Problems in risk analysis often involve extreme events, so direct empirical data for risk analysis are almost always lacking. One way to address this dilemma is to borrow strength from "outside" of the direct data. This paper first analyzes the possibilities of using existing methods for strength borrowing. Inspired by Hierarchical Overlapping Coordination (HOC) and based on HBM, this paper further proposes a new modeling method, Hierarchical Coordinated Bayesian Model (HCBM). In HCBM, multiple non-nested hierarchies for strength borrowing can be constructed from multiple perspectives, and then the multiple hierarchies will be coordinated under a generalized linear model framework. With HCBM, information can be borrowed from multiple dimensions. Case studies are carried out to that HCBM is a very powerful tool to address the data scarce problem in risk analysis.

An Environmental Sampling Model for Combining Judgment and Randomly Placed Samples

❖ Landon Sego, Pacific Northwest National Laboratory, P0 Box 999, Richland, WA 99352, *landon.sego@pnl.gov*; Anderson Kevin, Pacific Northwest National Laboratory; Brett Matzke, Pacific Northwest National Laboratory; Brent Pulsipher, Pacific Northwest National Laboratory; William K. Sieber, National Institute for Occupational Safety and Health; Stanley Shulman, National Institute for Occupational Safety and Health; Matt Gillen, National Institute for Occupational Safety and Health; James Bennett, National Institute for Occupational Safety and Health; Wilson John, Pacific Northwest National Laboratory; Richard Gilbert, Pacific Northwest National Laboratory

Key Words: environmental sampling, judgment sampling, Bayesian modeling, anthrax, emergency response, bioterrorism

In the event of the release of a lethal agent (such as anthrax) inside a building, law enforcement and public health responders take samples to identify and characterize the contamination. Sample locations may be rapidly chosen based on available incident details and professional judgment. To achieve greater (and quantifiable) confidence of whether or not a room or zone was contaminated, or to certify that detectable contamination is not present after decontamination, it is desirable to augment judgment samples with samples that are taken from randomly selected locations. We consider a Bayesian model for combining the information gained from both judgment and randomly placed samples. We investigated the sensitivity of the model to the parameter inputs and we also discuss various scenarios under which the model may be employed.

Unveiling Significant Patterns in High-Speed Networks

❖ Jin Cao, Bell Labs, 3 Ross Ave, Edison, NJ 08820, *cao@research.bell-labs.com*; Tian Bu, Bell Labs; Aiyou Chen, Alcatel-Lucent; Patrick Lee, Columbia University

Key Words: monitoring, sequential hashing, linear regression, sparse signal

In many real time monitoring applications, the identification of anomalous elements is of great importance. Often, the number of elements N to be monitored are humongous, but number of anomalous elements H is much fewer. Define the anomalous behavior by a quantity that is of high volume. We propose a novel algorithm that requires only $O(H\log N)$ both in memory and computational overhead to uncover the anomalous elements, instead of $O(N)$ by simply tracking all elements individually. Our algorithm couples a sequential hashing scheme for encoding the original data and a linear regression procedure for estimating the values of anomalous elements. Using examples in high speed network monitoring, we demonstrate using both theoretical analysis and real Internet traces studies that our approach can achieve very high accuracy while maintaining a low memory and computation overhead.

425 Modeling and Evaluating Customer Behavior

Section on Statistics and Marketing

Wednesday, August 1, 10:30 am–12:20 pm

A Bayesian Approach for the Multidimensional Scaling of Ordered Preference Data

❖Duncan K.H. Fong, The Pennsylvania State University, 456 Business Building, Smeal College of Business, University Park, PA 16802, *i2v@psu.edu*

Key Words: Multidimensional Scaling, Bayesian Analysis, Preference Analysis

We develop a Bayesian MDS vector model to analyze ordered preference data. Unlike classical MDS procedures, the Bayesian method includes a probability based criterion to determine the number of dimensions of the derived joint space map. Also, our procedure models the raw data which ameliorates the need of any data preprocessing as required for some widely used classical MDS procedures. A unique feature of the proposed Bayesian procedure is that it allows external attribute information to be directly incorporated into the spatial representation of the preference data, and that the derived prior eliminates the need of identification constraints on the bilinear structure of the model to obtain a proper posterior distribution. We solve the computational problem associated with the use of an intractable posterior distribution from one data set as a prior for the analysis of another dataset.

Quantifying the Indirect Effects of a Marketing Contact

❖ Dingxi Qiu, Northwestern University, 2145 Sheridan Road, Room C210, Evanston, IL 60208, *d-qiu@northwestern.edu*; Edward C. Malthouse, Northwestern University; Ajit Tamhane, Northwestern University

Key Words: Marginal Effect, Survial Analysis

We quantify the indirect effects of marketing contact points (e.g., exposure to a banner ad or catalog) on outcomes such as the likelihood or amount of purchase. A direct effect of a banner ad is a purchase resulting from a click on the ad while an indirect effect is the incremental change in the likelihood of purchase at another time due to the exposure. An indirect effect of a catalog is a purchase from a store or web site. Models that do not account for indirect effects underestimate the value of the contact. We develop survival analysis models of the marginal effects of such contacts, enabling companies to quantify the benefit of exposing someone to an ad, which can be compared with costs. The approach is tested with real data in three contexts: banner ads for an on-line service, catalogs from a multi-channel retailer, and donation solicitations from not-for-profit organizations.

Development of Personal Proposals from Financial Institutes During Interactive Real-Time Browsing of Potential Clients

❖ Hye-Kyoung Kim, Intellidyn Corp., 105 Maxess Road, Suite 121, Melville, NY 11747, *dkuznetsov@intellidyn.com*; Dmitri Kuznetsov, Intellidyn Corp.; Peter E. Harvey, Intellidyn Corp.

Key Words: financial services, interactive browsing, marketing

Internet browsing presents significant opportunities to challenge attention of potential clients to financial and other services that fit their needs in the best possible way. At present there exist many segmentation and clustering schemes that describe behavioral peculiarities and needs of any person belonging to one of these segments/clusters. Thus, to make a reasonable proposal we have to assign a browsing person to a specific cluster. The problem is how we can do it during a brief real-time connection with (1) a fun, (2) just a few unostentatious questions, and (3) a confidential, respectful and clear manner. In the present paper we described our approach to the problem that was successfully applied at a leading financial-services company.

Channel Partner Runoff Detection for Retention

❖ Wei Chai, GE Global Research, One Research circle, Niskayuna, NY 12309, *chai@research.ge.com*; Deniz Senturk-Doganaksoy, GE Global Research; Kareem Aggour, GE Global Research; James Williams, GE Global Research

Key Words: retention, run off, dormancy, life cycle, heatmap, detection

By reselling goods and services, partners provide the primary source of end customers for many businesses. Retaining them is therefore critical for a business to remain vital over time. A†process and online tool has been developed†to detect†and display statistically significant changes in the volume of customers provided by each partner, enabling sales and marketing to accurately identify retention issues early. This is accomplished through three steps: 1) modeling the typical partner lifecycle for†the business, 2) creating a detection algorithm to classify each partner into one of four stages (growing, maintaining, declining or dormant) over time, and 3) visualizing the key metrics and partner states. In just the first month of use, this enabled the identification of†over 1k declining or†newly†dormant partners, resulting in sales calls leading to over $4MM in new volume submitted.

Dynamics of Opinion Spreading in Marketing

❖ Dmitri Kuznetsov, Intellidyn Corp., 175 Derby street, Unit 40, Hingham, MA 02043, *dmitrikuznetsov@yahoo.com*

Key Words: Sociophysics, mediaphysics, word of mouth, opinion spreading, marketing

Word of mouth and opinion spreading are key points of many social phenomena including marketing and politics. At present only sociophysics proposed methods that allow analyzing opinion spreading deep enough and disclosing the corresponding causality. However, majority of the sociophysical approaches are far from business realities because of inability to operate with large data sets. In 2005 we introduced Mediaphysics that adopted advantages of both sociophysics and statistics with ability to treat real-life data. This approach uncovered very challenging and complicated peculiarities of marketing and political voting in the presence of word of mouth, including response dynamics of social systems to reach a stationary state which can take days or even months. These aspects of marketing represent the subject of this paper and a part of the corresponding methodologies at Intellidyn Corp.

On Models for Identifying Likely Lapsers

❖ Enping Mai, Syracuse University, 311 Whitman School of Management, 721 University Ave, Syracuse, NY 13244, *emai@syr.edu*; Raja Velu, Syracuse University

Key Words: Lapse Models, Inference, Random effects

Identifying likely lapsers is very important to almost all firms. The two models that are proposed in the marketing literature to predict customer lifetime duration are the Pareto/NBD model by Schmittlein et al (1987) and the BG/NBD model by Fader et al (2005). They mainly differ how the dropout rates are modeled. While those papers consider only estimation of the models, in our research we extend their models. In addition we develop formal inference procedures, further incorporating both time varying covariates such as promotions, and variables that are fixed over time, such as demographic information. The results show that these can vastly improve the performance of the models.

Evaluating Database Marketing Models: More Than Meets the Eye

❖Samuel Koslowsky, Harte Hanks Analytics, 55 Fifth Avenue, New York, NY 10003, *sam_koslowsky@harte-hanks.com*

Key Words: database marketing, response models, gains

Unlike ten or so years ago, most managers would agree that database marketing statistical models provide significant advantages. While this, in itself, is certainly an encouraging sign, model evaluation is not as simple a matter as many would believe. Many ingredients of the evaluation process are overlooked. Some believe that with the presence of a gains or performance report, the evaluation criteria are obvious. The standard frequently summoned is "does the decile report seem reasonable?" While many managers and analysts would walk away happy with viewing a monotonically decreasing gains report, there are additional dimensions that can impact model performance, but are frequently ignored by both marketers and data miners. These include percent captured variations, choppiness, stability, predictability, and explainability.

426 Correlated Data Methods and Analyses in Epidemiology ●

Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences, Section on Health Policy Statistics

Wednesday, August 1, 10:30 am–12:20 pm

Analysis of Smoking Cessation Patterns Using a Stochastic Mixed Effects Model with a Latent Cured State

❖Sheng Luo, Johns Hopkins University, 615 N Wolfe St, Dept. of Biostatistics, E3038, Baltimore, MD 21205, *sluo@yahoo.com*; Ciprian M. Crainiceanu, Johns Hopkins University; Thomas A. Louis, Johns Hopkins University; Nilanjan Chatterjee, National Cancer Institute

Key Words: Cure Model, Mixed-effect Model, Recurrent Events, Stochastic Transition Model, Smoking Cessation

We develop a mixed model to capture the complex transition processes among addiction and non-addiction stages of tobacco dependence. An important innovation of our model is allowing an unobserved cure state, or permanent quitting, in contrast to transient quitting. This distinction is necessary when censoring prevents unambiguous determination of being "cured." We apply our methodology to a large (29,133 subjects) longitudinal study to model smoking cessation patterns using a discrete-time stochastic mixed-effect model with three states: smoking, transient cessation and permanent cessation. Random subject specific transition probabilities among these states are used to account for subject-to-subject heterogeneity. Another innovation is to design computationally practical methods using the marginal likelihood obtained by integrating over the Beta distribution of random effects.

Assessing Baseline and Post-Baseline Effects in Repeated Measures Data Analyses: An Example Using Smoking Cessation Data

❖ Xiaolei Zhou, RTI International, 3040 Cornwallis Road, P O Box 12194, Durham, NC 27703, *xzhou@rti.org*; Beth Sherrill, RTI International; Jianmin Wang, RTI Health Solutions

Key Words: repeated measures, time interaction, time-dependent variables

In behavior modification studies dynamically changing factors have complex effects on subsequent actions. The question is how to accommodate baseline and post-baseline factors in one flexible model and assess effects on outcomes that are of clinical relevance. For example, in a longitudinal study about smoking relapse, potential predictors for successful smoking cessation include: nicotine dependence, mood disturbance, and weight change. Several models could be fit to the data depending on the clinical question of interest; often in exploratory analyses, however, several models are fit to the data. This presents a particular challenge to the analyst if competing inferences are made. We developed a single flexible model that can be used to test a variety of null hypotheses, thus providing the analyst a more efficient tool for evaluating complex time-varying covariates.

Systolic Blood Pressure in Childhood Predicts Hypertension and Metabolic Syndrome Later in Life

❖Shumei Sun, Wright State University, 3171 Research Blvd, Kettering, OH 45420, *shumei.sun@wright.edu*; Ruohong Liang, Wright State University

Key Words: Metabolic Syndrome, Obesity

The availability of long-term serial data from the Fels Longitudinal Study presents the opportunity to link hypertension and the metabolic syndrome in adulthood directly to blood pressures measured decades earlier in the same individuals as children and to establish criterion values for blood pressure that predict hypertension and the metabolic syndrome later in life. Serial data from 240 men and 253 women in the Fels Longitudinal Study were analyzed to derive age- and sex-specific childhood blood pressures that predict hypertension and the metabolic syndrome in adulthood using a random effects model in a discovery sample and validated these criterion values in a larger sample using logistic regression. Children with systolic blood pressures above the criterion values established in this longitudinal study are at increased risk of hypertension and the metabolic syndrome later in life.

Modeling Correlated Longitudinal Processes with Some Processes Partially Observed

❖Qilu Yu, Johns Hopkins University, 2024 E Monument St, Baltimore, MD 21205, *qyu4@jhmi.edu*; Laurel Beckett, University of California, Davis; David Bennett, Rush University; Robert Wilson, Rush University

Key Words: longitudinal data, aging, partially observed process, profile likelihood, EM algorithm

Complex processes in growth and aging may have multiple markers of change, and result from a complex interplay between risk factors. Our goal is to model correlated processes and to understand their relationship over time. Such studies encounter particular challenges when some processes are only partially observed. The motivating example is pathology-mediated clinical change in aging, where the pathology development can only be observed with a single post-mortem measurement. We propose a family of mixed effects models for investigating the pathologic mechanisms linking risk factors to clinical decline. According to different model assumptions, two sets of approaches are developed. In Method I we propose a profile likelihood method, while in Method II a generalized EM algorithm is adapted for parameter estimation. We illustrate the analysis with an example from the Religious Orders Study.

Assessing Trend in Longitudinal Data

❖Shailendra Banerjee, Centers for Disease Control and Prevention, 5375 Country Lake Court, Lilburn, GA 30047, *snb1@cdc.gov*

Key Words: trend, Longitudinal data, Regression, Autocorrelation

In a longitudinal data analysis, researcher is often confronted with the task of explaining the trend in the response variable. A simple regression model with time as the independent variable may be affected by errors caused by variables not represented in the model. The autocorrelated residuals violate the basic assumptions of linear regression analysis. If the error variances are not properly estimated, very important aspects of the regression analysis will be incorrect. Although the estimates of the parameters would not be biased, assessments regarding their variances would be incorrect, and the significance tests and confidence intervals of the parameters would be in error. In this paper, several alternative methods will be discussed to reduce the autocorrelation and test their effect on trend evaluation. Other methods of estimating trend will also be compared.

Regression Analysis of Multivariate Panel Count Data

❖ Xin He, University of Missouri-Columbia, 1548 Towne Dr Apt 3E, Columbia, MO 65202, *xhw4c@mizzou.edu*; Xingwei Tong, Beijing Normal University; Jianguo Sun, University of Missouri-Columbia; Richard J. Cook, University of Waterloo

Key Words: Counting processes, Estimating equations, Marginal mean model, Multivariate recurrent events, Observation processes

Panel count data frequently occur in periodic follow-up studies that concern recurrence rates of some recurrent events. Fields that produce such data include epidemiological studies, medical follow-up studies, reliability studies and tumorigenicity experiments. Multivariate panel count data arise if more than one type of recurrent events are of interest and this article discusses regression analysis of multivariate panel count data. For inference, we present a class of marginal mean models which leave the dependence structures for related types of recurrent events completely unspecified. Some estimating equations are developed for inference and the resulting estimates are shown to be consistent and asymptotically normal. Simulation studies are conducted for practical situations and the methodology is applied to a motivating cohort study of patients with psoriatic arthritis.

Analysis of Recurrent Event Data in Case-Crossover Studies

❖ Xianghua Luo, The University of Minnesota, 420 Delaware St SE, MMC 303, Minneapolis, MN 55455, *luox0054@umn.edu*; Chiung-Yu Huang, National Institutes of Health; Gary S. Sorock, Johns Hopkins Bloomberg School of Public Health

Key Words: conditional logistic regression, incidence rate ratio, Mantel-Haenszel, within-subject averaging

The case-crossover design is useful for studying the transient effects of exposures on risk of events when only data on cases are available. While the original design was proposed for univariate failure time data, in many applications recurrent events (e.g., elderly falls and school absences) are often encountered. In such situations, the within-subject dependence among recurrent events needs to be taken into account. Existing approaches rely heavily on independence assumptions for the recurrent events, either through a non-testable Poisson process assumption or a hypothetically complete set of observed covariates. We propose to adjust the Mantel-Haenszel estimator of incidence rate ratio and the conditional logistic regression, using a within-in cluster averaging technique. The proposed methods are applied to a study of the effect of medication changes on falls among the elderly.

427 Topics in Confidence Intervals and Inference ●

Section on Survey Research Methods

Wednesday, August 1, 10:30 am–12:20 pm

Parametric Bootstrap Confidence Interval in a Small-Area Estimation Problem

❖ Huilin Li, University of Maryland, Rm. 4315 Math Building, College Park, MD 20740, *huilin@math.umd.edu*; Partha Lahiri, University of Maryland

Key Words: Parametric Bootstrap, confidence interval, small area estimation, Fay-Herriot model, ADM estimator

In this paper, we developed the theory of using ADM variance estimator in a parametric bootstrap method in constructing confidence intervals of small area means. Using a Monte Carlo simulation study, we first investigate the performance of different variance component methods in the parametric bootstrap confidence interval, and then we compare our method with different rival methods in terms of coverage probabilities and average lengths. We then demonstrate the utility of the parametric bootstrap method by analyzing several real life datasets.

Setting Precision Requirements for Estimating Proportions

❖ James R. Chromy, RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709, *jrc@rti.org*

Key Words: Effective sample size, Suppression

For proportions, we know how to prescribe the effective sample size required if the required estimate, its expected values, and the precision requirements are provided. After the data are analyzed, you may be asked if some estimates should be suppressed because they lack precision. Both problems require reasonable criteria which, for binary data, vary based on the true population values. Even if direct estimates are used, the logistic transformation may be used to construct asymmetric confidence intervals on the proportion scale which are bounded between 0 and 1. This paper explores criteria related to standard errors, relative standard errors, and standard errors on the logit scale. It is assumed that the survey will involve multiple estimates with a range of values of the true proportion.

Confidence Intervals for Statistics for Categorical Variables from Complex Samples

❖ Valerie Hsu, Westat, 1650 Research Blvd, Rockville, MD 20850, *valeriehsu@westat.com*; Keith Rust, Westat

Key Words: Asymptotically Normal Estimators, Multi-Stage Sampling, Log Transformation, Log(-Log) Transformation, Index of Disparity, WesVar

The standard large-sample-based approach to deriving a two-sided confidence interval performs poorly for proportions that are close to zero or one. Alternative methods are suggested in the literature, but no one approach appears to be universally superior, and some of the approaches are difficult to implement in practice, especially when the data are collected via a complex sample design. Such problems are even more evident when one estimates certain summary measures for two-way tables, such as the Index of Disparity (Pearcy and Keppel, 2002), where the parameter estimated is defined to be non-negative, but estimates close to zero are common. In this paper we consider the practical implementation of some alternatives to deriving confidence intervals in these cases using statistical software for the analysis of complex survey data (WesVar).

One-Sided Coverage Intervals for the Extreme Binomial Proportion Under Complex Sample Design

❖ Yan Liu, Internal Revenue Service, Statistics of Income, P O Box 2608, Washington, DC 20013, *yan.k.liu@irs.gov*; Phillip S. Kott, USDA National Agricultural Statistics Service

Key Words: binomial proportion, continuity correction, coverage interval, coverage probability, effective sample size, stratified sample design

The interval estimation of a binomial proportion is difficult, especially when the proportion is extreme (very small or very large) compared to the sample size. Most of the methods proposed in the literature implicitly assume simple random sampling. These interval-estimation methods are not immediately applicable to data derived from a complex sample design. Some recent papers have addressed this problem, proposing modifications for complex samples. Matters are further complicated when a one-sided coverage interval is desired. This paper provides an extensive review of methods for constructing coverage intervals of a binomial proportion under both simple random and complex sample designs. It also evaluates the empirical performances of different methods for constructing one-sided coverage intervals for an extreme proportion under stratified simple random sampling.

Investigation of Empirical Bayesian Confidence Intervals for Expenditure Data in the Medical Expenditure Panel Survey

❖ Robert Baskin, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *robert.baskin@ahrq.hhs.gov*; Malay Ghosh, University of Florida

Key Words: gamma distribution, log-normal distribution, complex survey

The Medical Expenditure Panel Survey (MEPS), a national probability sample survey sponsored by the Agency for Healthcare Research and Quality (AHRQ), is designed to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. Research has shown the expenditure data from MEPS to have a marked positive skewness. Previous work has considered using different distribution functions for the confidence intervals of mean expenditures. The current project will evaluate the use of empirical Bayes confidence intervals for subdomain means using MEPS data. A simulation is used to compare the results.

Tailoring Survey Data Products and Reliability Criteria To Best Match Customers' Revenue Needs

❖ Mandy Webb, Arbitron Inc., 9705 Patuxent Woods Drive, Columbia, MD 21046, *mandy.webb@arbitron.com*; Alan R. Tupek, Arbitron Inc.

Key Words: Margin or Error, Sample Size, Reliability Criteria

Arbitron produces listening estimates used by radio stations and advertisers alike to buy and sell advertising time. Some market estimates are based on relatively small sample sizes, which should affect how clients use these ratings in making purchasing decisions. This paper details a project by which Arbitron attempts to work with and explain to clients the effect of sample size on margin of error. Emphasis is placed on how to use margin of errors to better understand the reliability of small domain estimates as well as developing best practices for the users of our data. Alternative client solutions are explored such as two-book averaging and other options for increasing sample sizes at the market level.

Properties of the Combined Ratio Estimator and a Best Linear Unbiased Estimator When Design Control Is Problematic

❖ Stephen Woodruff, Retired, 800 West View Terrace, Alexandria, VA 22301 2750, *stevewoodruff@comcast.net*

Key Words: Combined Ratio Estimator, Self-Weighting Design, Best Linear Unbiased Estimation, Stratified Cluster Designs

This is a study of a Best Linear Unbiased Estimator (BLUE) and the Combined Ratio Horwitz-Thompson Estimator (CRHT) for sampling problems where efficient self-weighting designs are impractical. Sampling of mail is an example of this. A measure for "self-weighting" is derived and their variances as functions of this measure are compared under repeated sampling from stratified cluster designs. For self-weighting designs, the variances of CRHT and BLUE are similar but that of CRHT increases rapidly as the measure increases. The BLUE is insensitive to the measure. When design control is impossible, the variance of CRHT can be many times (2 to 100s) that of the BLUE. The BLUE is derived from mail population characteristics and the sample design, making model failure a virtual non-issue. Simulation results, supporting mathematics, their implications, and applications are presented.

428 Analysis and Disclosure of Survey Data

Section on Survey Research Methods

Wednesday, August 1, 10:30 am–12:20 pm

Multiyear Estimation for a Single Stratum from the Continuous National Health and Nutrition Examination Survey (NHANES)

❖ R. Lee Harding, Westat, 1650 Research Blvd, Rockville, MD 20850, *leeharding@westat.com*; Jay Clark, Westat; Jill M. Montaquila, Westat

Key Words: variance estimation, combined independent samples

NHANES is a continuous household survey with nationally representative annual samples. Design-based estimates of rates of health-related outcomes for demographic domains were produced for one major stratum by combining the data across six years of data collection. However, estimating the variance of these domain estimates was problematic because of modifications to the sample design within the major stratum over the period: the number of primary sampling units (PSUs) varied between 19 and 23 across years; the substratification scheme varied; and PSU measures of size were updated. This paper discusses the issues involved in producing estimates and variance estimates for this stratum, and describes the methodology used. Consideration is also given to the production of estimates for other strata, where the numbers of PSUs may be smaller.

The Effectiveness of Stepwise Discriminate Analysis as a Post Hoc Procedure to a Significant Manova

❖ Erik Heiny, Western Carolina University, Department of Mathematics and CS, Cullowhee, NC 28723, *eheiny@wcu.edu*

Key Words: stepwise, discriminant, analysis, manova

This study examined stepwise discriminant analysis (DA) as a post hoc procedure to a significant MANOVA. Monte Carlo simulations were conducted under the assumptions of normality and homogeneity, using different levels of sample size, effect size, a, number of dependent variables, and correlation structure among the dependent variables. High correlations among dependent variables and large numbers of dependent variables led to poor results for power and/or type I error using stepwise DA. However, under certain conditions, stepwise DA performed quite well. The most important factor is keeping the number of dependent variables small, in this study no more than 3. As values for p and sample size increased, type I error became too difficult to control, even for small a, and therefore stepwise DA is not recommended under these conditions.

Researcher Access to Microdata: Measuring the Impact of Different Access Protocols on Data Utility

❖Stephanie Shipp, National Institute of Standards and Technology, 100 Bureau Drive, Bldg 101 Rm A301, Gaithersburg, MD 20899-4710, *stephanie.shipp@nist.gov*; Julia Lane, NORC at the University of Chicago

Key Words: confidentiality, data access, data quality

Ensuring researcher access to data in a manner that simultaneously protects the privacy and confidentiality of human subjects and permits high quality research is a major challenge. A major innovation in this paper is that it describes a set of access protocols that combines elements from the computing and social sciences to develop secure remote data access protocols that provide technical security, as well as mechanisms to ensure confidence in computing and storage systems. The paper addresses three research questions. By how much can technical tools reduce the risk of disclosure associated with data access? What technical procedures have the least impact on research quality, conditional on their cost? What incentives are effective in promoting data documentation and use?

Measuring Privacy Loss Due to Data Release: A New Perspective

❖Lei Liu, Yale University, 300 George St Suite 523, New Haven, CT 06511, *lei.liu@yale.edu*; Heping Zhang, Yale University

Key Words: privacy, disclosure control, confidentiality, mathematical model

A basic concern of data collectors releasing data to third parties is that the privacy of individuals in the data set might be compromised. When data releasing becomes necessary, the quantification of privacy loss is very important for data administrators since the design of the data releasing scheme must consult such measures to control the privacy loss. In this paper, we try to construct privacy loss measures from the viewpoints of data providers and keepers. The new measures can be defined without modeling intruders' behaviors and provide an unified rational for many existing privacy loss measures under various situations in the literature.

The Use of Propensity Score Matching Methods in Mixed Mode Studies

❖John Bremer, Harris Interactive, 161 Sixth Avenue, New York, NY 10013, *jbremer@harrisinteractive.com*; Kerri Miller, Harris Interactive

Key Words: statistical matching methods, multi-mode surveys, survey research, propensity score, MSE, bias

The use of multiple data collection modes in a single survey instrument is gaining in popularity. Different modes allow for more flexible control of the sample and for more advanced methodologies to be used to investigate the problem at hand. Little has been written with regard to how to combine data from different data collection modes. Often the data from different modes are combined in a simple manner by simple stacking. This technique can produce bias in the results. This paper proposes the use of an application of propensity score matching in order to more accurately combine data from multiple sources. There is a tradeoff between bias and variability in incorporating this technique but generically the MSE of using a propensity score approach is lower than the MSE of using a stacking method. This paper will examine the conditions under which this is specifically the case.

Practical Considerations in Applying the pq-Rule for Primary Disclosure Suppressions

❖Meghan O'Malley, Bureau of Labor Statistics, 2 Massachusetts Ave, Washington, DC 20212, *meghan.s.omalley@gmail.com*; Lawrence Ernst, Bureau of Labor Statistics

Key Words: Confidentiality, p-Percent Rule, pq-Rule, Sample Weights, Imputation, Primary Suppressions

As statistical agencies try to "move as far as possible toward the use of a small number of standardized disclosure limitation methods whose effectiveness has been demonstrated" (Working Paper 22), many practical difficulties arise in applying standard sensitivity measures to particular data sets. This paper discusses possible ways of handling several common complications which arise when applying the pq-rule. These include analysis of changes in disclosure risk due to imputation, suggestions for use of weights as protection, and discussion of handling final weights less than one (generally arising from controlling sampled statistics to independent universe values). Ideas in this paper come mainly from considering how the pq-Rule could be applied to the Occupational Employment Statistics survey conducted by the U.S. Bureau of Labor Statistics in partnership with the States.

429 Section on Health Policy Statistics Speaker with Lunch (fee event)

Section on Health Policy Statistics

Wednesday, August 1, 12:30 pm–1:50 pm

Importance of Using Cost-Effectiveness Analysis To Inform Health Policy in the United States

❖David Meltzer, The University of Chicago, 5841 S Maryland, Chicago, IL 60637, *dmeltzer@medicine.bsd.uchicago.edu*

Key Words: public policy, cost-effective analysis

Medical cost-effectiveness analysis (CEA) can provide information on the value of medical technologies relative to their cost, and has been widely used by policy makers in many countries to improve the efficiency of health care spending. CEA has not yet played a major role in shaping medical coverage decisions in the United States, but the pending crisis in the financial viability of Medicare and recent establishment of the Medicare prescription drug benefit have created the opportunity for CEA to play an important role in U.S. policy. One barrier to increasing the role of CEA in U.S. health policy is concerns about the validity of methods. This talk will illustrate the application of CEA methods to address resource allocation problems in health care and discuss key methodological challenges that need to be addressed as CEA moves from theory to practice in the United States.

430 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section
Wednesday, August 1, 12:30 pm–1:50 pm

Incomplete Data in Longitudinal Studies

❖Geert Molenberghs, Hasselt University, Center for Statistics, Agoralaan 1, Diepenbeek, B3590 Belgium, *geert.molenberghs@uhasselt.be*

Key Words: incomplete data, missing data, clinical trials, ignorable analysis

Most, if not all, clinical and nonclinical longitudinal studies are subject to incompleteness. In a biopharmaceutical context, a lot of relatively simple methods for dealing with such incomplete data have been in use. On the other hand, ever more advanced methods for handling incomplete longitudinal studies are proposed by the research community. Regulatory authority-based, biopharmaceutical, and academic researchers need to work together to bridge this gap and move on. What would be viable strategies to achieve this goal?

Clinical Trial Simulation: Beyond Statistics

❖Bret Musser, Merck & Co., Inc., 126 E Lincoln Ave, RY34A316, Rahway, NJ 07083, *bret_musser@merck.com*

Key Words: clinical trials, simulation, decision making

Can statistical modeling and simulation improve decisionmaking in drug development? As we industrialize our approach to clinical trials, how should we—statisticians—be involved in bringing order to the chaos of how to design clinical development programs?

Statistical Concerns and Methodology Regarding Radiographic Data in Rheumatoid Arthritis

❖Lifen Zhou, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, *lifenz@amgen.com*

Radiographic data are commonly used to assess structural joint damage in RA. The use of statistics to analyze this data is important and often needs additional consideration due to the nature of radiographic disease progression. Some key characteristics of this data will be highlighted. Statistical methodology in this area will be discussed, including the standard rank method, such as Wilcoxon rank; covariates adjusted method, such as Quade; and the strata variables adjusted method, such as Van Elteren.

Design-Based Multiplicity Adjusting Strategies in Late-Phase Clinical Trials

❖Jianjun Li, Merck & Co., Inc., 351 N Sumenytown Pike, North Wales, 19454, *jianjun_li@merck.com*

Key Words: late-phase clinical trials, primary endpoints, secondary endpoints, PAAS, Hochberg procedure, gate-keeping procedures

In clinical trials, endpoints are classified as primary, secondary, etc. To control trial-wise, type-I error, the Hochberg procedure is commonly used for handling multiple testing among primary endpoints. For secondary endpoints which can be used for regulatory claims but are potentially under-powered, the PAAS has been proposed. For secondary endpoints, which play only a supportive role, gate-keeping procedures have been developed. While these procedures are proposed for general use, they do not incorporate a unique characteristic of late-phase clinical trials: The primary endpoints are powered by design. This discussion will focus on how to utilize this important property in analysis of multiple endpoints. Recent developments in statistical methodology and their impacts on clinical trials such as CAPRICORN trial, ATLAS study, DEFINITE trial, and PROactive study, will be included.

431 Business and Economics Statistics Section Roundtables with Lunch (fee event)

Business and Economics Statistics Section
Wednesday, August 1, 12:30 pm–1:50 pm

Model Selection

❖David Findley, U.S. Census Bureau, DIR, Room 5K152B, Washington, DC 20233, *david.f.findley@census.gov*

Key Words: CP, AIC, BIC, GIC, MDL, Bootstrap

Since Mallows, Akaike, Schwarz, and Rissanen introduced their model selection criteria, there has been steady growth in the related literature, developing variants of their now classical criteria, as well as alternative criteria, for situations involving small samples, robust or suboptimal estimation methods, semiparametric models, mis-specified models, etc. We can talk about practical model selection problems and these newer criteria.

Statistical Analysis of Large Panels of Economic Data

❖Serena Ng, University of Michigan, Dept. of Economics, 317 Lorch Hall, Ann Arbor, MI 48109, *serena.ng@umich.edu*

Key Words: principal components, boosting, lasso, regularization, forecasting, factor models

More economic data are available for analysis as we move forward in calendar time and as technological progress makes it possible to collect information about more series. Developing statistical methods that can efficiently exploit the large volume of information available is therefore important. In this session, we will discuss statistical methods for analyzing large panels of economic data. While principal components is often used to reduce the dimension of economic time series, methods such as regularization (e.g., LASSO) and adaptive learning (e.g., boosting) are more popular in machine learning and gene classification. We will discuss the strengths and weaknesses of these methods with emphasis on their role in economic forecasting and financial analysis.

432 Section on Bayesian Statistical Science Roundtable with Lunch (fee event)

Section on Bayesian Statistical Science
Wednesday, August 1, 12:30 pm–1:50 pm

The Quasi-Bayesian Approach to Risk Management at the Department of Homeland Security

❖David Banks, Duke University, ISDS Box 90251, Durham, NC 27708, *banks@stat.duke.edu*

Key Words: game theory, risk, DHS

The DHS uses a somewhat Bayesian approach to risk management. We will outline what is currently done, and then open the floor for creative conversation. Some of the issues that arise include risk analysis, subjective probability, the Kadane-Larkey approach to game theory, and portfolio theory for maximizing return on defensive investments.

433 Section on Physical and Engineering Sciences Roundtables with Lunch (fee event)

Section on Physical and Engineering Sciences
Wednesday, August 1, 12:30 pm–1:50 pm

Verifying Computer Experiments with Physical Experiments

❖William Notz, The Ohio State University, Dept. of Statistics, 1958 Neil Avenue, Columbus, OH 43210-1247, *win@stat.ohio-state.edu*

Key Words: computer experiments, validation

Complex computer simulations have become popular substitutes for physical processes that are expensive or difficult to observe. Experimenting with the computer simulation as a means of studying the physical process is called a computer experiment. For this approach to yield meaningful results, the computer simulation must be a reasonable approximation to the physical experiment. Verifying this is sometimes referred to as validating the computer experiment. This session will consider methods for validating computer experiments. There is only a modest literature on validating computer experiments. We will discuss the current state of the literature on this topic, problems that are related (such as calibration and combining data from multiple sources), and possible research directions. A bibliography of some of the literature that exists will be provided.

Publishing an Industrial Case Study: Advice for Would-Be Authors

❖Robert Rodriguez, SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-8000, *bob.rodriguez@sas.com*

Key Words: Case Studies, Industrial statistics

Good published case studies of industrial statistical applications are in high demand by practicing statisticians and classroom instructors, but they are in short supply. This is partly because many statisticians working in industry are not encouraged or prepared to write for journals. We will discuss success factors for publishing industrial case studies (such as an application of general interest and statistical substance) and how to approach the writing process. Join us if you have material you would like to publish, or if you are affiliated with a journal that would like to attract industrial case studies.

434 Section on Quality and Productivity Roundtables with Lunch (fee event)

Section on Quality and Productivity
Wednesday, August 1, 12:30 pm–1:50 pm

Communicating with Nonstatisticians

❖Mark Bailey, SAS Institute Inc., 215 Reillywood Avenue, Haddonfield, NJ 08033, *mark.bailey@sas.com*

Key Words: client, communication, presentation, collaboration

Discuss and share your ideas about the best ways to communicate statistical results to nonstatisticians in the context of the client-statistician relationship. This discussion will include experiences from successful and failed collaborations. In particular, we want to know what worked well and what didn't.

Risk-Adjusted Charts in Health-Care Monitoring

❖Landon Sego, Pacific Northwest National Laboratory, P0 Box 999, Richland, WA 99352, *landon.sego@pnl.gov*

Key Words: risk adjustment, control chart, cusum, survival time, health care, monitoring

Statistical process control (SPC) is making notable inroads in the health care industry. The application of SPC we will consider here involves monitoring the mortality associated with a health care procedure, which, in order to be effective, requires accounting for the fact that the risk of mortality for each patient depends not only on the quality of the procedure, but also on risk factors that are unique to each patient that exist prior to the procedure. Incorporating these risk factors (covariates) into the monitoring scheme is known as risk adjustment. We will discuss the risk-adjusted varieties of several well-known charts: Shewhart p-chart, the Sets method, the resetting SPRT, and the Bernoulli Cusum chart. We also will discuss a new risk-adjusted Cusum chart for monitoring right-censored survival times and consider areas of future development.

435 Section on Statistical Consulting Roundtable with Lunch (fee event)

Section on Statistical Consulting
Wednesday, August 1, 12:30 pm–1:50 pm

Recognizing and Resolving Breakdowns in Your Consulting Practice

❖Doug Zahn, Zahn and Associates, 249 Timberlane Road, Tallahassee, FL 32312, *zahn@stat.fsu.edu*

Key Words: breakdown, consulting, interpersonal, intrapersonal, technical, bless the breakdown

Breakdowns are a natural part of everyone's consulting practice. No matter how carefully we plan and communicate, breakdowns (failures of the conversation, project, process, etc., to function as intended) will inevitably occur in encounters with our clients, colleagues, supervisors, and staff. In this luncheon, I will introduce you to a process for recognizing and resolving your breakdowns. You will have an opportunity to use it to identify at least one step you can take to resolve your most vexing, repeater breakdown.

436 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education
Wednesday, August 1, 12:30 pm–1:50 pm

User-Friendly Jargon for Statistics

❖Larry Weldon, Simon Fraser University, Statistics and Actuarial Science, 8888 University Drive, Burnaby, BC V5A 1S6 Canada, *weldon@sfu.ca*

Key Words: parametric, nonparametric, data analysis, undergraduate, resampling, graphics

Significant results that are not important, normal distributions that are not usual, standard deviations that are not acceptable, regression that is not a backward step, expected values that never occur - these common occurrences have the potential to confuse all but the most indoctrinated statistician. And then we verbalize our descriptions of these things in a foreign language using Greek letters. Is there a fix to this disastrous marketing of our discipline?

Statistics and Research Design: Oil and Vinegar?

❖Seth Hirshorn, University of Michigan-Dearborn, D21 FCS 19000 Hubbard Drive, Dearborn, MI 48105, *shirsh@umich.edu*

Key Words: Proposals, Research Design, Data Analysis, Evaluation, Education, Statistics

The education of statisticians and researchers divides the subjects of data analysis and research methodology into separate seminars using texts dedicated to either statistics or methodology but not both. Textbooks are focused on one or the other but fail to integrate the two in ways that "work" in the real world of research and evaluation. The intersection of these is many times a "tipping point" for proposal reviewers, yet overlooked by proposal writers. In this context we will discuss how statistics and research data area successfully integrated in developing proposals. We will also consider how statistics and methodology seminars may be transformed resulting in an integration of these subjects. Members of the community with an interest and/or experience in such an "integrated approach" - case studies, especially, are invited and welcomed.

Academic Integrity in a High-Tech Classroom

❖John McKenzie, Babson College, MathScience Division, Babson Park, MA 02457-0310, *mckenzie@babson.edu*

Key Words: cheating, education, examinations, laptops, quizzes, technology

Technology is an integral component of any applied statistics course today. But, there is a dark side to the omnipresence of technology. With the increasing ability of students to communicate with others in and outside of the classroom, the possibility of academic integrity violations increases. In a wireless environment, this illegal exchange of information may be accomplished by using a laptop, PDA, or even a cell phone. And, while an instructor may ban PDAs or cell phones, it is difficult to restrict laptops when students are using them for calculations or note-taking. This roundtable will discuss ways to reduce such unauthorized collaboration when laptops are allowed in the classroom. Among the preventative measures to be discussed are blocking devices and electronic quizzes and examinations. If time permits, other examples of inappropriate use of technology will be considered.

437 Section on Statistical Graphics Roundtable with Lunch (fee event)

Section on Statistical Graphics
Wednesday, August 1, 12:30 pm–1:50 pm

Visualizing Model and Parameter Uncertainty

❖David W. Scott, Rice University, 4143 Marlowe Street, Houston, TX 77005-1953, *scottdw@rice.edu*

Key Words: Visualization, Parameter Estimation, Collinearity, Model Sensitivity

Understanding the uncertainties of model parameter estimates and corresponding model sensitivities with many parameters offers an excellent opportunity for visualization experimentation. We will begin with topics in regression and mixture models.

438 Section on Statisticians in Defense and National Security Roundtable with Lunch (fee event)

Section on Statisticians in Defense and National Security
Wednesday, August 1, 12:30 pm–1:50 pm

Text Data Mining for Defense Applications

❖Jeffrey Solka, Naval Surface Warfare Center, 17320 Dahlgren Road, Dahlgren, VA 22448-5100, *jeffrey.solka@navy.mil*

Key Words: text, data, mining, clustering, discovery, future

This roundtable discussion will explore the role of text data mining in defense applications. Discussions will include, but not be limited to, the role of text data mining in the characterization of country capabilities, its role in the characterization of the state of the art of a discipline area, and its role in discovery. Discussion will focus on the speaker's experiences in this area and his knowledge of the state of the text data mining literature. We also will explore who the customers might be for these techniques and where the future lies, both in the technology and in the important problems that have not yet been addressed.

439 Section on Statistics in Epidemiology Roundtable with Lunch (fee event)

Section on Statistics in Epidemiology
Wednesday, August 1, 12:30 pm–1:50 pm

Longitudinal Causal Modeling

❖James Robins, Harvard School of Public Health, 677 Huntington Ave, Suit 821, Boston, MA 02115, *robins@hsph.harvard.edu*

Key Words: causal inference, marginal structural model, structural nested model, confounding, semiparametric models

In longitudinal studies with time-varying treatment, the standard approach to the estimation of the effect of treatment on an outcome of interest is to model the outcome at time t as a function of past treatment history. This approach may be biased, whether or not one further adjusts for the past history of time-dependent confounding covariates, when these covariates predict subsequent outcome and treatment history and are, themselves, influenced by past treatment. In this roundtable, I will discuss several methods that can provide, under certain assumptions, valid estimates of the causal effect of a time varying treatment in the presence of time varying confounding factors. These include inverse-probability-of-treatment-weighted (IPTW) estimation of a marginal structural model, g-estimation of a structural nested model, and the parametric g-computation algorithm. The three methods will be compared.

440 Section on Survey Research Methods Roundtable with Lunch (fee event)

Section on Survey Research Methods
Wednesday, August 1, 12:30 pm–1:50 pm

Auditing Election Results

❖Wendy Rotz, Ernst & Young, 1225 Connecticut Ave, Washington, DC 20036, *wendy.rotz@ey.com*

Key Words: election, audit, vote, ballot, trail, poll

Voter confidence and transparent elections are essential to modern democracies. A statistical foundation in the development of election auditing procedures would aid transparency and improve voter confidence in election results and the democratic process in the United States and abroad. Statistical issues may include 100% recounts in close elections, random recounts in all elections, paper trails, Six Sigma applications from quality process control, exit polling, and many more. Share ideas concerning statistical methodologies for auditing election results.

441 Social Statistics Section Roundtable with Lunch (fee event)

Social Statistics Section
Wednesday, August 1, 12:30 pm–1:50 pm

Media Coverage Regarding Data on Mortality in Iraq

❖Leslie Roberts, Columbia Mailman School of Public Health, 60 Haven Ave., Level B-4, New York, NY 10032, *les@a-znet.com*; Gilbert Burnham, Johns Hopkins University; Shannon Doocy, Johns Hopkins University

Key Words: cluster sample, survey, Iraq, mortality

Between May and July 2006, a national mortality survey was conducted in Iraq. The results were reported in The Lancet in October of 2006; the survey found that by July 2006, there had been 655,000 (CI: 393,000–943,000) excess deaths above the pre-war baseline in Iraq as a consequence of the war. Media coverage of the results emphasized the controversial nature of the study, in spite of the conventional methods employed, and has continued to focus on lower surveillance estimates based on morgue and hospital records. This is because of the difficulties of communicating sample results, the extremely incomplete nature of vital record-keeping in settings of extreme violence, and successful spinning of the story by parties with nonscientific agendas. The discussion leader will outline his experiences with the press as an author of this study.

442 Statistical Society of Canada Roundtable with Lunch (fee event)

SSC
Wednesday, August 1, 12:30 pm–1:50 pm

Fifty Years of Empirical Bayes/Shrinkage Inference: Past, Present, and Future Outlook

❖S. Ejaz Ahmed, University of Windsor, Department of Mathematics and Statistics, 401 Sunset AV, Windsor, ON N9B 3P4 Canada, *seahmed@uwindsor.ca*

Key Words: Empirical Bayes, shrinkage estimation, high-dimensional data, likelihood inference

Empirical Bayes (EB) and shrinkage estimation methods surfaced in statistical literature about 50 years ago. Shrinkage and EB methods provide useful techniques for combining data from various sources. Very large gains in inferential accuracy may be achieved by judiciously employing these strategies. It seems inconceivable, then, that not much attention has been paid to these methods in the development of statistical theory. Current statistical problems in the high-dimensional and low-sample size settings have regenerated interest in these techniques. These methods continue to play vital roles in the area of statistical inference, and Professor Efron has predicted continued research through the early 21st century. The goal of this roundtable is to bring together frequentist and Bayesian perspectives and explore and create common ground.

443 Late-Breaking Session 2: War and Death in Iraq ●

The ASA, ENAR, SSC, WNAR, IMS, Section on Survey Research Methods, Section on Statisticians in Defense and National Security, Social Statistics Section
Wednesday, August 1, 2:00 pm–3:50 pm

Results of a Survey Estimating Mortality Before and After the 2003 Invasion of Iraq

❖Leslie Roberts, Columbia Mailman School of Public Health, 60 Haven Ave., Level B-4, New York, NY 10032, *les@a-znet.com*

Methods: Between May and July 2006, a national mortality survey was conducted in Iraq using 50 clusters randomly selected among 16 governorates proportionate to size (achieved 47) of 40 households (achieved 39). Information on deaths from 1,849 households containing 12,812 persons was collected. Findings: 1474 births and 629 deaths were reported during the observation period. Death rates were 5.5/1000/yr (CI: 4.3-7.10) pre-invasion, and overall post-invasion, 13.3/1000/yr (CI: 10.9-16.1) for the roughly 40 months post-invasion. The death rates through September 2004 closely paralleled the results of our 2004 survey. We estimate that through July 2006, there have been 655,000 (CI: 393,000-943,000) excess deaths in Iraq as a consequence of the war. Of post-invasion deaths, 601,000 (CI: 426,000-794,000) were due to violence.

444 Recent Advances in Causal Modeling ● ✪

ENAR, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Biometrics Section, Section on Teaching Statistics in the Health Sciences, WNAR

Wednesday, August 1, 2:00 pm–3:50 pm

Mediation Analysis Using Random Assignment Interacted with Baseline Covariates as Instrumental Variables

❖Dylan Small, University of Pennsylvania, 400 Huntsman Hall, 3730 Walnut St., Philadelphia, PA 19104, *dsmall@wharton.upenn.edu*

Key Words: instrumental variables, mediation analysis, causal inference

Randomized trials often involve multi-component treatments. Mediation analysis seeks to “open up” such “black box” treatments and explain their mechanism of action. Traditional approaches to mediation analysis (e.g., Baron and Kenny, 1986) make the strong assumption that mediating variables are randomized within the levels of the baseline randomization variable. Ten Have et al. (2005) introduced a structural mean model approach that makes potentially more tenable assumptions when a covariate is available that predicts the mediating variable. We show how Ten Have et al.’s assumptions can be interpreted in terms of instrumental variables (IVs) and present an IV approach that can be implemented in standard software. We develop a sensitivity analysis for our approach and apply our approach to a trial of an intervention for treating depression in primary care practice.

Inference in Randomly Censored Regression Models Using Conditional Moment Inequalities

❖Shakeeb Khan, Duke University, 213 Social Sciences Building, Durham, NC 27708, *shakeebk@duke.edu*

Key Words: conditional moment inequalities, quantile minimum distance, endogeneity

Under a conditional quantile restriction, randomly censored regression models can be written in terms of conditional moment inequalities. These inequalities restrict the parameters to a set. We then show regular point identification can be achieved under a set of interpretable sufficient conditions. Our results generalize existing work on randomly censored models in that we allow for covariate dependent censoring, endogenous censoring and endogenous regressors. Maintaining the point identification conditions, we propose a quantile minimum distance estimator which converges at the parametric rate and has an asymptotically normal distribution. A small scale simulation study and an application using drug relapse data demonstrate satisfactory finite sample performance, and the ability of our proposed method to control for selective compliance.

Causal Comparisons in Randomized Trials of Two Active Treatments

Jason Roy, Geisinger Health System; ❖Joseph W. Hogan, Brown University, Box G-S121, Providence, RI 02912, *jwh@brown.edu*

Key Words: noncompliance, smoking cessation, potential outcomes, bounds, principal stratification, Bayesian inference

In behavioral medicine trials, such as smoking cessation trials, two or more active treatments are often compared. Causal parameters of interest might include those defined by subpopulations based on their potential compliance status under each assignment. Principal stratification is an effective method for inference, but even when subjects in one arm do not have access to the other treatment, the causal effect of each treatment typically can only be identified within certain bounds. We propose to use additional information—compliance-predictive covariates—to improve identification of the causal effects. We demonstrate the benefit of utilizing covariate information in both a simulation study and in an analysis of data from a smoking cessation trial.

445 Twenty-Five Years of Noninferiority Trials ● ✪

Biopharmaceutical Section, ENAR, Biometrics Section, WNAR

Wednesday, August 1, 2:00 pm–3:50 pm

Early Developments in Clinical Equivalence and Similarity Trials

❖William C. Blackwelder, University of Maryland, Baltimore County, c/o 8613 Hempstead Avenue, Bethesda, MD 20817-6711, *wcb99@comcast.net*

Key Words: Clinical equivalence, Similarity, Noninferiority, Clinical trials

Studies designed to show a new treatment is not worse than an active control by as much as a specified quantity are now common. Work in the late 1970s on a study of reduced duration of therapy for serious fungal infection led to the 1982 Controlled Clinical Trials paper. In a 1978 paper Makuch and Simon had addressed the same problem. “Equivalence” and “similarity” have been used to describe this type of trial. “Noninferiority” is now generally used, but it must be understood as shorthand for “not inferior by as much as a specified quantity”. Null hypotheses of non-equality have also been used in other settings - for example, in trials designed to show vaccine efficacy is greater than a specified positive value.

The Establishment of Efficacy on the Basis of Noninferiority Trials: Updating Our Approach

❖Steven Snapinn, Amgen Inc., One Amgen Center Drive, 24-2-C, Thousand Oaks, CA 91320, *ssnapinn@amgen.com*

Key Words: Preservation of effect, Constancy assumption, Non-inferiority margin, Discounting, Synthesis Method

This presentation will address two aspects of the usual approach used to establish the efficacy of an experimental treatment on the basis of a non-inferiority trial. First, the concept of setting a non-inferiority margin is routinely used and is recommended in various regulatory guidance documents; however, it is extremely inefficient relative to an approach which pools data from the non-inferiority trial with historical data comparing the active control with placebo. Second, there is the concept that an experimental treatment must demonstrate preservation of the active control’s effect. While this appears on its surface to be a reasonable requirement, on closer inspection it can be shown to lead to serious logical inconsistencies. In this presentation I will propose a new paradigm for establishing efficacy that avoids the inconsistencies and inefficiencies of our current approach.

Bayesian Perspectives on the Issue of Noninferiority

❖Andrew P. Grieve, King’s College, Department of Public Health Sciences, School of Medicine, Floor 7 Capital House, 42 Weston St, London, SE1 3QD United Kingdom, *Andy.Grieve@kcl.ac.uk*

Key Words: Non-Inferiority, Bayesian methods, Posterior probabailities

Since the publication of Bill Blackwelder’s initial paper on non-inferiority there has been an increasing interest in the topic as is evidenced by the volume of publications. At the same time there has been a re-emergence of interest in Bayesian methods in Pharmaceutical R&D as has recently been recognized by initiatives of both US (Critical Path Initiative) and European

(Innovative Medicines Initiative) regulatory authorities. In this paper we investigate the intersection of these two topics and ask the question does a Bayesian perspective bring additional insight into the realm of noninferiority assessment.

446 Recent Advances in Statistical Learning: Modern Regularization Methods and Semisupervised Learning ● ✪

Section on Statistical Computing, Section on Nonparametric Statistics

Wednesday, August 1, 2:00 pm–3:50 pm

Efficient Large-Margin Semisupervised Learning

Junhui Wang, Columbia University; ❖Xiaotong Shen, The University of Minnesota, School of Statistics, Minneapolis, MN 55455, *xshen@stat.umn.edu*

Key Words: Margin, Semisupservised, Generalization

In classification, semi-supervised learning involves a large amount of unlabeled data with only a small number of labeled data. This imposes great challenge in that the class probability given input can not be well estimated through labeled data alone. This talk presents a large margin semi-supervised learning method that constructs an efficient loss to measure the contribution of unlabeled instances to classification. An iterative scheme is derived for implementation. The method is examined with two large margin classifiers: support vector machines and psi-learning. Our theoretical and numerical analyses indicate that the method achieves the desired objective of delivering high performance.

Infinite Dimensional Lasso

❖Nathan Srebro, Toyota Technological Institute at Chicago, 1427 East 60th Street, Chicago, IL 60637, *natisrebro@gmail.com; nati@uchicago.edu*; Saharon Rosset, IBM T.J. Watson Research Center; Ji Zhu, University of Michigan; Grzegorz Swirszcz, IBM T.J. Watson Research Center

Key Words: Lasso, Kernel methods, Sparsity

We describe a practical methodology for fitting ℓ_1 regularized prediction models in very high, or even infinite, dimensional feature spaces. It is based on extensions of path-following methods for ℓ_1 regularized models for the lasso, 1-norm SVM and other modeling problems. We show that the sparsity property of Lasso holds in infinite dimensions as well, and discuss learning performance. We illustrate our approach on the problem of fitting additive regression splines subject to a total variation penalty (equivalent to ℓ_1 penalty). The resulting method offers great flexibility in fitting additive models to data, and we demonstrate its performance on one simulation example and two real-life datasets.

Grouped and Hierarchical Model Selection through Composite Absolute Penalties

Peng Zhao, University of California, Berkeley; ❖Guilherme V. Rocha, University of California, Berkeley, 367 Evans Hall, Department of Statistics, Berkeley, CA 94720, *gvrocha@stat.berkeley.edu*; Bin Yu, University of California, Berkeley

Key Words: regularization, penalization, regression, classification, variable selection

For datasets with many predictors and few samples, side information often must be added to fitting. We introduce Composite Absolute Penalties (CAP) to blend predefined grouping and hierarchical information among the predictors into regression and classification. Special cases include Zou & Hastie(2005)'s elastic net, Kim et al.(2005)'s Blockwise Sparse Regression and Yuan & Lin(2006)'s GLASSO. CAPs are built by combining norm penalties at the across and within group levels. For disjoint groups, a Bayesian interpretation lays bare the role of the norms used to construct CAP. Hierarchical selection is reached by defining nested groups. For general CAPs, we use the BLASSO and cross-validation to compute CAP estimates. For CAPs built from L_1 and L_\infty norms, we give efficient algorithms and regularization selection criteria. The feasibility of CAP is shown through simulated experiments.

447 Risk Analysis in National Defense ●

Section on Risk Analysis, Section on Statisticians in Defense and National Security, Section on Physical and Engineering Sciences

Wednesday, August 1, 2:00 pm–3:50 pm

Risk Estimation and National Security: Introduction and Overview

❖A. John Bailer, Miami University, Dept. of Math. & Stat., Oxford, OH 45056, *baileraj@muohio.edu*

Key Words: hazards, exposure, dose-response models, uncertainty

Risk estimation might be used to set exposure limits to chemicals in the workplace, to determine allowable contaminant levels for wastewater treatment plants or to establish engineering targets for operating characteristics of the space shuttle. Other examples include the risk of failure of safety systems in nuclear power plants or buildings in response to seismic events. Concerns for the risk of terrorism are not far from conscious thoughts in days of color-coded travel alerts. These concerns reflect a low probability-high impact hazard. In this overview, an introduction to risk assessment ideas including hazard identification, exposure assessment, dose-response modeling and risk characterization will be presented. The distinction between risk assessment for chemical agents and failure of complex systems will be discussed and risk of terrorism will be placed in this context.

Biosurveillance and the BioSense Program

❖Henry Rolka, Centers for Disease Control and Prevention, 1600 Clifton Road, MS E06, Atlanta, GA 30333, *HRolka@CDC.Gov*

Key Words: biosurveillance, surveillance, situational awareness

Empirical biosurveillance systems frequently use temporal and geographic contexts to establish a baseline against which to compare recent data. Numerous varieties of data types are used in varying levels of geographic scope across levels of public health. There are many components all of which must develop and function in concert in order for success in early event detection or situational awareness. A brief background and history of what has been done in this area at CDC will be presented with an update of the BioSense Program.

Risk Analysis at Los Alamos National Laboratory

❖Alyson G. Wilson, Los Alamos National Laboratory, P.O. Box 1663, MS F600, Los Alamos, NM 87545, *agw@lanl.gov*

Key Words: risk analysis, national security, defense, complex systems

Risk analysis at Los Alamos focuses around the analysis of complex systems. This talk focuses around the steps that we use to develop analyses: representing the systems, mapping the data, analyzing the data, and making decisions about future data collection. These are the steps used for risk assessment and risk management. Each step is illustrated using examples from the Department of Defense, Department of Energy, or Department of Homeland Security.

448 Preventing, Correcting, and Measuring Overcoverage in Censuses ● ✪

SSC

Wednesday, August 1, 2:00 pm–3:50 pm

Preventing and Correcting Coverage Errors in the United States Decennial Census

❖David Sheppard, U.S. Census Bureau, 8023 Delores Court, Chesapeake Beach, MD 20732-4546, *david.w.sheppard@census.gov*

Key Words: Decennial, followup, US Census Bureau

Ensuring that every person in the United States is counted once and only once is a vital goal of the Decennial Census. For many decades, the Census Bureau has evaluated coverage in the Census and documented that people are missed, counted more than once, or counted in the wrong place. In response, we have researched ways to help respondents initially report the appropriate household members, as well as looked for ways to identify households that did not. One of the primary means we use to correct the household roster when it is not enumerated correctly is through a coverage follow-up interview. In this paper, we present some of the results of the 2005 National Census Test and discuss the direction of the coverage follow-up operation for the 2010 Decennial Census.

Unduplication in the 2010 United States Decennial Census

❖Robin A. Pennington, U.S. Census Bureau, 4600 Silver Hill Road, Mail Stop 7600, Washington, DC 20233, *robin.a.pennington@census.gov*

Key Words: overcoverage

In Census 2000, the U.S. Census Bureau began using Optical Character Recognition to capture the names on questionnaires. As a result, it is now possible to check for duplicate people after the questionnaires are processed. However, there are challenges to removing duplicates from final Census counts. Duplicated persons occur as a result of: persons listed on more than one questionnaire for reasons such as part-time residency, and housing frame duplication. In the former, it is not known from the questionnaires which residence is correct. In the case of housing duplication, resolution is complicated by the similar situation of duplication of persons resulting from apartment mix-ups. The Census 2010 plan is for a general coverage operation following enumeration to resolve identified coverage cases. In this paper we present an overview of our unduplication plans for the 2010 census.

Measuring Person Duplication: The 2006 Canadian Census Overcoverage Study

❖Heather Farr, Statistics Canada, 15-C R.H. Coats building, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *heather.farr@statcan.ca*; José Morel, Statistics Canada

Key Words: record linkage, coverage error

This paper describes a two step methodology for measuring person duplication in a census. Firstly, an exact match is conducted between census records and an auxiliary file of administrative records covering most of the Census target population to find easily identified duplicates, based on name and demographic variables. A name frequency file is used to assign first and last names. Secondly, an estimate of less easily identified duplicates is derived using clerical verification of a sample of potential duplicates obtained by probabilistic record linkage, based on name and demographic variables. Geographical proximity of addresses is also considered when calculating the strength of the probabilistic linkage. The methodology, its advantages over previous methods used at Statistics Canada, innovative features, and initial quantitative evaluations of the exact match methodology are outlined.

Overcoverage in the 2011 UK Census

❖Owen Abbott, Office for National Statistics, Room 4200N, ONS, Segensworth Road, Titchfield, Fareham, International PO155RR United Kingdom, *owen.abbott@ons.gov.uk*; James Brown, University of Southampton

Key Words: 2011 Census, census coverage, overcoverage

For the 2001 UK Census, the Office for National Statistics (ONS) instigated the One Number Census (ONC) project which focused on measuring undercoverage, since overcoverage had historically not been a significant issue in previous censuses. Evidence gathered during the ONC and subsequent studies indicated that overcoverage in the 2001 Census may have been around 0.5%. For the 2011 UK Census it is expected that this figure will be larger due to changes in collection methodology and the rapid change in society and living arrangements. ONS is exploring ways for measuring overcoverage alongside the research to develop its methodology for measuring undercoverage. This presentation lays out how ONS is tackling this problem and the areas of research that will be addressed. Early ideas will be discussed, such as survey and matching based methods.

449 Sparse Inference and Large-scale Multiple Comparisons ● ✪

IMS, Section on Nonparametric Statistics

Wednesday, August 1, 2:00 pm–3:50 pm

Maxima of Discretely Sampled Random Fields, with an Application to 'Bubbles'

❖Keith J. Worsley, McGill University, Montréal, QC H3A 2K6 Canada, *keith.worsley@mcgill.ca*; Jonathan E. Taylor, University of Montréal

Key Words: Random fields, Improved Bonferroni inequalities, Bubbles, Euler characteristic

A smooth Gaussian random field with zero mean and unit variance is sampled on a discrete lattice, and we are interested in the exceedance probability or P-value of the maximum in a finite region. If the random field is smooth relative to the mesh size, then the P-value can be well approximated by results for the continuously sampled smooth random field (Adler, 1981; Worsley, 1995; Taylor & Adler, 2006). If the random field is not smooth, so that adjacent lattice values are nearly independent, then the usual Bonferroni bound is very accurate. The purpose of this talk is to bridge the gap between the two, and derive a simple, accurate upper bound for intermediate mesh sizes. The result uses a new improved Bonferroni-type bound based on discrete local maxima. We give an application to the "bubbles" technique for detecting areas of the face used to discriminate fear from happiness.

Network Kriging

❖Eric D. Kolaczyk, Boston University, Department of Mathematics and Statistics, 111 Cummington Street, Boston, MA 02215, *kolaczyk@math.bu.edu*; David Chua, Boston University; Mark Crovella, Boston University

Key Words: computer network, linear prediction

We consider the problem of monitoring certain path-based properties, such as loss rates or packet delays, across an entire computer network, based on a limited number of measurements on some subset of all possible paths. This problem is naturally formulated as one of statistical prediction, and we offer a simple class of predictors for standard quantities of interest, resulting in a methodology for `network kriging.' Linear algebraic methods of subset selection may be used to make effective choice of which paths to measure. The mean square prediction error properties of our overall approach are characterized through appropriate bounds. The methodology is illustrated through a handful of empirical examples. Crucial to the success of our approach is the low effective rank of so-called `routing matrices,' as observed in practice, which effectively induces a sparse inference problem.

Bayes and Empirical Bayes Approaches to Controlling the FDR

❖Cun-Hui Zhang, Rutgers University, Department of Statistics, Hill Center, Busch Campus, Piscataway, NJ 08854, *czhang@stat.rutgers.edu*; Weihua Tang, Rutgers University

Key Words: Bayes rule, empirical Bayes, false discovery rate, dependent data, multiple testing, Fourier method

We formulate a Bayes optimization problem as the maximization of the total amount of statistical discovery subject to a preassigned level of certain conditional false discovery rate, and propose an empirical Bayes approach based on the Bayes rule. The Bayes and thus the empirical Bayes approaches are formulated for general dependent data. The asymptotic optimality of the Benjamini-Hochberg rule is proved in the empirical Bayes sense. A Fourier method is used to estimate the proportion of true null hypotheses. A time series model is studied as an example of dependent data. Some simulation results are presented.

450 JASA, Applications and Case Studies Invited Session ● ✪

JASA, Applications and Case Studies, General Methodology

Wednesday, August 1, 2:00 pm–3:50 pm

Subjective Likelihood for the Assessment of Trends in the Ocean's Mixed-Layer Depth

❖Michael Lavine, Duke University, Box 90251, Durham, NC 27708, *michael@stat.duke.edu*; Ana Rappold, Duke University; Susan Lozier, Duke University

Key Words: subjective Bayes, physical oceanography, climate change, likelihood

This paper describes a Bayesian statistical analysis of long term changes in the depth of the ocean's mixed layer. The data are thermal profiles recorded by ships. For this data there is no good sampling model and therefore no obvious likelihood function. Our approach is to elicit posterior distributions for training data directly from the expert. We then infer the likelihood function and use it on large datasets.

451 Increasing Statistical Literacy in Nonstatistical Professions ✪

Committee on Outreach Education, Section on Statisticians in Defense and National Security

Wednesday, August 1, 2:00 pm–3:50 pm

Teaching Medical Statistics

❖Mary A. Foulkes, George Washington University, 6110 Executive Blvd., Suite 750, Rockville, MD 20852-3903; *mfoulkes@jhsph.edu*

Key Words: teaching, medical statistics, journalists, physicians

In order to contribute to the rigorous development, conduct, dissemination and interpretation of clinical research, both physicians and journalists need an understanding of the basic principles of clinical research design, and beyond. They need to appreciate the differences between the range of experimental and observational study designs, and the appropriate inferences from them. Those funding and conducting clinical research will rely on clinical researchers and the medical journalists who widely publicize results to provide sufficient information to inform public health policy. Training of both the physicians designing and conducting the research and the journalists reporting on results strongly impacts the effective, efficient and ethical conduct of both research and public policy. Experiences in teaching courses for oncologists, pediatric neurologists, etc. will be highlighted.

Lessons Learned in Teaching a Statistics Workshop

❖Jeffrey Solka, Naval Surface Warfare Center, 17320 Dahlgren Road, Dahlgren, VA 22448-5100, *jeffrey.solka@navy.mil*; Daniel Carr, George Mason University

Key Words: course, instruction, teaching, science, laymen

This talk details our recent experiences in conducting a workshop in "Communicating Statistics to the Uninitiated" at the American Association for the Advancement of Science. The talk will briefly discuss the material covered at the talk, the audience participation in the workshop, and perhaps most importantly the audience feedback. In addition some ideas to improve for future workshops on this topic will be presented.

Statistics in a Legal Context

❖Mary W. Gray, American University, 4400 Massachusetts Avenue NW, Dept of Mathematics and Statistics, Washington, DC 20016-8050, *mgray@american.edu*

Key Words: application to legal cases, expert witness, outreach to lawyers

Those in the legal profession frequently have difficulty understanding some of the most basic statistical concepts. The ASA outreach course is not intended to turn lawyers into statisticians but rather to give them the ability to deal with the fundamentals of data analysis in order to work effectively with expert witnesses. Descriptive statistics, t-tests, nonparametric tests, and regression are introduced and applications to legal cases are discussed. Cautions for both lawyers and statisticians as well as experiences in dealing with lawyers and judges are highlighted.

452 Adaptive Monte Carlo Methods

Section on Bayesian Statistical Science, ENAR

Wednesday, August 1, 2:00 pm–3:50 pm

Adaptive Independent Metropolis-Hastings Using Normal Mixtures

❖Robert Kohn, University of New South Wales, Sydney 2052 , NSW, Sydney, International 2052 Australia, *r.kohn@unsw.edu.au*; Paolo Giordani, Bank of Sweden

Key Words: Clustering, Markov chain Monte Carlo, Semiparametric regression models, State space models.

Adaptive Metropolis-Hastings samplers use information obtained from previous draws to tune the proposal distribution. The tuning is carried out automatically, often repeatedly, and continues after the burn-in period. We propose a method for adaptive independent Metropolis-Hastings sampling using a mixture of normals as a proposal distribution. To take full advantage of the potential of adaptive sampling our algorithm updates the mixture of normals frequently, starting early in the chain. The algorithm is built for speed and reliability and its sampling performance is evaluated with simulated examples and with applications to time-varying-parameter, semiparametric, and stochastic volatility models.

Combining Adaptation with MCMC Sampling Methods

❖Heikki Haario, Lappeenranta University of Technology, Skinnarilankatu 34, Lappeenranta, P.O.Box 20 Finland, *haario@users.csc.fi*; Antti Solonen, Lappeenranta University of Technology

Key Words: MCMC, adaptation

Several adaptive sampling strategies have been proved to be ergodic and tested with various benchmark examples. However, applications in real modeling cases may still provide challenges. Typical sources for the difficulties are high dimensionality, strongly correlated or non–identifiable parameters, bad initial proposal distribution or bad initial values for the sampling parameters. As a results, the adaptation might get properly started only after some hand tuning, while the very idea of adaptive schemes should be to make sampling more automatic. Here we show how adaptation may be combined with other approaches—Delayed Rejection, Reversible Jump MCMC, sampling from Gaussian mixtures—to overcome such difficulties. The real-life modeling examples come from chemical kinetics, biology, and remote sensing

Adaptive MCMC, Intractable Likelihood Functions, and Phylogenetic Inference

❖Yves F. Atchade, University of Michigan, Ann Arbor, MI 48105, *yvesa@umich.edu*

In this talk I will present a number of statistical models where a straightforward Bayesian approach is not possible due to the intractability of the likelihood function. A number of useful MCMC strategies to sample from the posterior distributions from such models will be discussed. An application to phylogenetic inference will be presented.

Adaptive Population Monte Carlo

❖Christian P. Robert, Universite Paris Dauphine, Place du Marechal de Lattre de Tassigny, Paris cedex 16, International 75775 France, *xian@ceremade.dauphine.fr*

Key Words: importance sampling, MCMC, clustering, tempering, entropy distance, tail exploration

Cappe et al. (2004) proposed a simulation scheme dubbed Population Monte Carlo which may be viewed as an Iterated Importance Sampling approach, with resampling and Markovian instrumental simulations. This scheme also is a particular case of the Sequential Monte Carlo Sampling approach of Delmoral et al. (2006), although we focus on the case where the target distribution is held fixed and the importance kernels are adapted during the iterations in order to optimize a performance criterion. In Douc et al. (2006), we established convergence properties of specific adaptation schemes in terms of entropy and variance improvements. The talk will cover recent developments based on preliminary clustering and tempering steps, obtained jointly with J.M. Marin and A. Mira, as well as applications to cosmological issues.

Recent Advances in Monte Carlo Methods

❖Christophe Andrieu, University of Bristol, University Walk, Bristol, BS8 1TW United Kingdom, *c.andrieu@bris.ac.uk*

Key Words: adaptive Monte Carlo methods, expected auxiliary variable method, sequential Monte Carlo methods

New challenges arising in particular from biology have stimulated the development of new types of Monte Carlo methods. Following [Haario et al. 1998 and Andrieu & Robert 2001], there has been a regain of interest in adaptive Monte Carlo methods, while we have recently developed in [Andrieu, Doucet & Roberts 2006] a new paradigm for Monte Carlo simulations called the “Expected Auxiliary Variable” approach, which covers and generalizes methods suggested to address some of the aforementioned challenges. I will show in this talk how both approaches can be combined to lead to efficient simulation algorithms.

453 Medallion Lecture V

IMS, General Methodology, Biometrics Section

Wednesday, August 1, 2:00 pm–3:50 pm

Joint Modeling of Longitudinal and Survival Data

❖Jane-Ling Wang, University of California, Davis, 1 Shields Avenue, Department of Statistics, Davis, CA 95616, *wang@wald.ucdavis.edu*

Key Words: semiparametric models, EM-algorithm, Monte Carlo integrations, functional data, informative dropout, measurement errors

It has become increasingly common to observe the survival time of a subject along with baseline and longitudinal covariates. Due to several complications, traditional approaches to marginally model the survival or longitudinal data encounter difficulties. Jointly modeling these two types of data emerges as an effective way to overcome these difficulties. We will discuss the challenges in this area and provide several solutions. One of the difficulties is with the likelihood approaches when the survival component is modeled semiparametrically as in Cox or accelerated failure time models. Several alternatives will be illustrated, including nonparametric MLEs, the method of sieves, and pseudo-likelihood approaches. Another difficulty has to do with the parametric modeling of the longitudinal component. Nonparametric alternatives will be considered to deal with this complication.

454 Multi-Criteria Prioritization and Ranking with Partial Order Without Compositing Multiple Indicators into an Index in Social, Environmental, and Infrastructure Work ● ✪

Environmental and Ecological Statistics, Section on Statistics and the Environment

Wednesday, August 1, 2:00 pm–3:50 pm

Exploring Partial Order of European Countries

❖ Paola Annoni, University of Milan, via Conservatorio 7, Dept. of Economics, Business and Statistics, Milan, 20122 Italy, *paola.annoni@unimi.it*; Rainer Bruggemann, Leibniz Institute of Freshwater Ecology and Inland Fisheries

Key Words: Hasse Diagrams, Dimension Theory in Posets, Planar Lattices, Formal Concept Analysis, Public Services

Partial Order Theory has been recently more and more employed to overcome the intrinsic disadvantage hidden in linear ranking, if a multiple indicator system is available. Despite its numerous positive features, there are cases where partial order interpretation can be troublesome. In these cases the analysis of underlying dimensions can uncover particular data structures. The paper shows a way of addressing the problem with the help of a case study, which deals with European opinions on public services. In particular, an overall ranking of countries is firstly provided and then a method to detect dimensions is discussed and applied. The analysis stems directly from the Partially Order Set and Lattice theory with particular references to dimension theory and Formal Concept Analysis. The study is eventually able to pinpoint role of services and criteria in defining the partial order.

Multicriteria Prioritization, Partial Order, and Hasse Diagrams in Environmental Sciences and Chemistry

❖ Rainer Bruggemann, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Muggelseedamm 310, Berlin, D-12587 Germany, *brg_home@web.de*

Key Words: Partial Order, Water Management, Chemicals

We take the point of view of environmetrics: Let first the data speak, and then let us include subjective preferences in order to get a unique decision. A set of partial orders (which themselves are partially ordered) is obtained. The iterative procedure stops if either a linear order is found, or two objects of specific interest are comparable. Two examples are given: As first example we evaluate the effects of nine water management strategies on the surface water system in Berlin. Surprisingly a conventional strategy can be recommended although the set of strategies also contained some high-technology strategies. The second example comes from the risk assessment of chemicals. It is shown that there are regions in the decision space, where it is not necessary to know the exact values of weights.

Unmasking Weight Camouflage of a Composite Index Based on Multiple Indicators

❖ K. Sham Bhat, The Pennsylvania State University, Center for Statistical Ecology, Department of Stat, 421 Thomas Building, University Park, PA 16802, *kgb130@psu.edu*; Ganapati P. Patil, The Pennsylvania State University

Key Words: Multicriterion prioritization, stepwise indicator aggregation, comparability acquisition profile, indicator consensus

Considerable work has been done on determining a method to accomplish an ordering of objects with multiple evaluation criteria. A weighted index is used to combine scores of multiple indicators as a criterion for ranking. Our goal is to understand the effect of the indicators and the weighting schemes on the ranking. We use two data analysis methods, POSAC and METEOR, to determine data based weights and to better understand the meaning behind the weights. We then compare data based and investigator based weighting schemes through their comparability acquisition profiles. The concepts of indicator consensus and indicator influence add additional insight. Poset theory and Hassediagrams are a basis for our analysis. These ideas are illustrated with four live data sets, each with their own challenges. Finally, we investigate the concepts of representability and measurement error.

455 Web Panel Surveys: A Solution to Declining Response Rates? ● ✪

Section on Survey Research Methods

Wednesday, August 1, 2:00 pm–3:50 pm

Nonresponse Bias in Two Methods of Panel Recruitment

❖ Charles DiSogra, Knowledge Networks, Inc., 1350 Willow Road, Suite 102, Menlo Park, CA 94025, *cdisogra@knowledgenetworks.com*; Daniel Slotwiner, Knowledge Networks, Inc.; Sarah Clinton, Knowledge Networks, Inc.; Elisa Chan, Knowledge Networks, Inc.; Erlina Hendarwan, Knowledge Networks, Inc.

Key Words: Panel, Non-response, bias, weighting

Knowledge Networks (KN) uses RDD sampling to recruit a national representative panel of 40,000+ members. Panelists participate in 3-4 surveys per month via Internet or WebTV (provided to non-Internet households). KN also conducts cross-sectional RDD surveys for some clients. One such survey offers an opportunity to join the KN panel at the end of a 20-minute interview about Yellow Pages use. This paper compares this post-interview recruiting to KN's usual method of RDD surveys done exclusively for recruiting. Several stages exist in the process of getting "survey-ready" paneled members, with attrition at each stage. This paper examines nonresponse at each stage for these two recruitment methods. Weighting adjustments are discussed for ameliorating non-response bias across several demographic dimensions to create a merged, nationally representative panel from both recruitment sources.

Combining RDD and Web Panel Surveys

❖ Karol Krotki, RTI International, 701 13th St NW, Suite 750, Washington, DC 20005, *kkrotki@rti.org*

Key Words: Coverage, Nonresponse bias, Survey mode

In light of decreasing response rates, it is important to research alternative modes. One such approach is based on web panel surveys that are relatively inexpensive and may become more viable as the proportion of Internet households increases beyond the current 70%. We discuss a 2005 smoking survey in Ohio that was implemented both via RDD and web panel. We first describe the weighting adjustments that were used to bring the data sets in line with population benchmarks. We then identify differences with respect to basic underlying demographics as well as substantive variables related to smoking. We explore models to identify differences and propose methods to adjusting for these differences. We conclude that web panel data can be made to resemble corresponding RDD survey data and that there are robust methods for combining data from these two survey modes.

Sample Matching for Web Surveys: Theory and Application

❖ Douglas Rivers, Stanford University, Dept of Political Science, 315 Encina West, Stanford, CA 94305, *rivers@stanford.edu*

Key Words: Web survey, panel, matching

Methods are described for selecting a sample from a pool of respondents with unknown selection mechanism when a suitable frame (or random sample with known selection probabilities) is available which contains auxiliary information that can be matched to the pool of respondents. Asymptotic properties of estimators are discussed and empirical examples using an opt-in Web panel are provided.

456 Statistical Graphics for Analysis of Drug Safety and Efficacy ●

Section on Statistical Graphics, Section on Risk Analysis, WNAR

Wednesday, August 1, 2:00 pm–3:50 pm

Statistical and Graphical Analysis of Adverse Event Counts in Clinical Trials

❖ Michael O'Connell, Insightful Corporation, 5317 Highgate Dr, Durham, NC 27713, *moconnell@insightful.com*

Assessment of drug safety is a primary goal in drug development. However, in contrast to the design and analysis of efficacy endpoints, there is typically little planning or innovative statistical analysis of safety data. In particular, adverse events are usually reported as tables of counts for treatment and control, sometimes accompanied by results from a statistical test that is not well-suited to the data structure. In this talk we present some statistical and graphical analyses of adverse event data that address issues such as the sparse nature of the count data and the many adverse event terms under consideration. Statistical models considered include hierarchical generalized linear models, least angle regression and ensemble methods such as random forests. Results from these analyses are presented as S-PLUSÆ graphical summaries that highlight aspects of drug risks.

Using Graphics To Discover and Explore

❖ Julia Wang, Johnson & Johnson PRD, 920 Route 202 S, PO Box 300, Raritan, NJ 08869, *jwang@prdus.jnj.com*

Key Words: Statistical graphics, S-plus

Rapid advancements have been made during the last 15 years or so in statistical visualization techniques. Statistical software such as S-plus provides an extensive set of tools to construct graphics to aid in the interpretation and presentation of clinical data. In this presentation, several examples will be given to illustrate the importance of graphing subject-level data and the limitations of traditional summary statistics-based graphics.

Statistical Graphics for the Analysis of Safety and Efficacy Data from Clinical Trials

❖ Ohad Amit, GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, PA 19426, *ohad.amit@gsk.com*

Key Words: Graphics, Clinical trial, efficacy, safety

Thorough evaluation of safety and efficacy has always been a key focus in the development of new pharmaceutical products. The predominant method for statistical evaluation and interpretation of safety and efficacy data collected in a clinical trial is a tabular display of descriptive and inferential statistics. There is a great opportunity to enhance evaluation of clinical trial data through the use of graphical displays, which can convey multiple pieces of information concisely and more effectively than can tables. Graphs can be used in an exploratory setting to help identify emerging safety signals, or in a confirmatory setting as a tool to elucidate the safety and efficacy of pharmaceutical product. We developed several graphical displays for routine data collected during a clinical trial, covering a broad range of graphical techniques.

Graphical Analyses of Clinical Trial Safety Data

Kefei Zhou, Amgen Inc.; ❖ Haijun Ma, Amgen Inc., One Amgen Ceter Dr, Mail Stop 24 2A, Thousand Oaks, CA 91320, *hma@amgen.com*; Hong A. Xia, Amgen Inc.; Matthew Austin, Amgen Inc.; George Li, Amgen Inc.; Michael O'Connell, Insightful Corporation

Key Words: safety analysis, adverse events, graphics, clinical trial safety, Insightful Miner, safety signal detection

Safety assessment is crucial in drug development. Pharmaceutical companies and regulatory agencies collect large amounts of safety data, which usually include adverse events, laboratory measurements, concomitant medications, etc. However, the analysis of safety data often has not received sufficient attention. As part of risk management, safety data should be continuously monitored and analyzed during clinical trials. Graphical analyses can maximize the ability to detect any unexpected and unusual features in clinical trial safety data. We present several graphs that have been developed to better visualize different types of clinical trial safety data and facilitate safety signal detection. These graphs are conveniently embedded in the Insightful Miner workflow, which not only allows users to analyze the data more interactively, but also streamlines the safety monitoring process.

Design of Statistical Graphics for Clinical Data

❖ Richard M. Heiberger, Temple University, Department of Statistics, Speakman Hall, Philadelphia, PA 19122-6083, *rmh@temple.edu*

Key Words: statistical graphics, safety data, adverse effects

Graphical displays, which can convey multiple pieces of information concisely and more effectively than can tables, are valuable in the identification and analysis of safety issues. We developed several graphical displays for routine safety data collected during a clinical trial, covering a broad range of graphical techniques. We discuss in detail the statistical and graphical principles underlying the production and interpretation of the displays.

457 Phase I Trials in Cancer ● ✪

Biometrics Section, WNAR, Biopharmaceutical Section, ENAR

Wednesday, August 1, 2:00 pm–3:50 pm

New Phase I Trial Designs for Combinations of Two Agents

Shenghua (Kelly) Fan, California State University, East Bay; Alan P. Venook, University of California, San Francisco; ❖ Ying Lu, University of California, San Francisco, Department of Radiology, Comprehensive Cancer Center, San Francisco, CA 94143-0946, *ying.lu@radiology.ucsf.edu*

Key Words: Phase I trial, MTD, DLT, toxicity, isotonic, dose escalation

Combination therapy of two anti-tumor agents may provide treatment synergy. Phase I trials search the maximum tolerated dose (MTD) of such combinations. The conventional approach is to pre-select an escalation path, usually increasing dose of one agent and then another, and to use the standard 3+3 procedure. This procedure may miss the optimum dose combination not on the path and prolong the duration to reach the MTD. In this paper, we present a comprehensive search for a two-agent combination therapy. We evaluate algorithms based on dose limiting toxicity (DLT) and combination of DLT and any toxicity in such a search. A two-dimensional isotonic estimation method for toxicity rate will be provided. We use simulation methods to compare 2+1 versus 3+3 algorithms. We conclude that a comprehensive search for MTD for a combination-therapy of two-agents can be practical and efficient.

An Application of the Time-to-Event Continual Reassessment Method: Dose Escalation in Patients with Lymphoma Undergoing Chemotherapy over an Extended Period

❖ Ken Cheung, Columbia University, 722 West 168th Street Rm 641, New York, NY 10032, *yc632@columbia.edu*

Key Words: Late toxicity, Cancer

Traditional phase I designs require each patient or small group of patients to be completely followed before the next patient or group is enrolled. In situations where chemotherapy is given over an extended period, this may result in trials of impractically long duration. To evade the timeliness issue in practice, the safety endpoint is often defined with respect to a short time period (28 days): it will underestimate the harm if toxicity tends to appear beyond the 28-day window. In the talk I first review the time-to-event continual reassessment method (TITE-CRM), which allows patients to be entered in a staggered fashion while utilizing observations up to the full follow-up period from each patient. In the second half, I summarize the outcomes of a recent lymphoma trial that adopts the TITE-CRM, and evaluate in retrospect the relative merits of the TITE-CRM and the standard method.

A Hierarchical Bayesian Design for Phase I Trials of Novel Drug Combinations

❖ Thomas M. Braun, University of Michigan, 1420 Washington Heights, M4063 SPH II, Ann Arbor, MI 48109, *nuarbmot@yahoo.com*

Key Words: Bayesian statistics, Adaptive designs, Phase I clinical trial, Drug discovery

Phase I clinical trials are increasingly designed to examine dual-agent therapies, but existing Phase I clinical trial designs identify the MTD of a single agent, and few published designs for Phase I trials of two agents exist. We propose a hierarchical Bayesian design for Phase I clinical trials designed specifically to examine the combined toxicity of two previously studied agents. We directly model the combined effects of both agents on the probability of toxicity and constrain the model so that the probability of toxicity increases with dose increases of either agent. We use adaptive Bayesian methods to determine the combination assigned to each subject. Through simulations, we demonstrate excellent operating characteristics for our design, in terms of identifying the optimal combination, as well as assigning a sufficient proportion of study subjects to the optimal combination.

An Adaptive Bayesian Approach to Incorporating PG Information in Phase I Trials

❖ Meihua Wang, University of Pittsburgh, Department of Biostatistics, Pittsburgh, 15260, *wang_m@nsabp.pitt.edu*; Roger Day, University of Pittsburgh

Key Words: Phase I trials, adaptive Bayesian design, genetic variations, PK, PD

Genetic variations can lead to huge inter-individual differences in drug efficacy and toxicity, so it is desirable to individualize chemotherapy. Genetic variations that affect drug metabolism are expected to affect all pharmacodynamic (PD) processes downstream, including both efficacy and toxicity, while other genetic variations may affect only one PD process. This suggests that a hierarchical model could help in designing an improved Phase I trial incorporating specific genetic information and prior belief. On the other hand, Phase I trials are small and may not contain enough information to support a highly adaptive design. The question is whether such a design could be helpful. We propose an adaptive Bayesian design based on a three-level hierarchical PK/PD model, incorporating patient response, toxicity, genotype and prior belief. Preliminary simulation results are encouraging.

Toward a Benchmark for Sequential Phase I Cancer Trial Designs

❖ Jay Bartroff, University of Southern California, Department of Mathematics, 3620 South Vermont Ave, KAP 108, Los Angeles, CA 90089-2532, *bartroff@usc.edu*; Tze Leung Lai, Stanford University

Key Words: phase I, sequential design, dynamic programming, Monte Carlo, dose finding

Phase I cancer trials have two competing aims: treating advanced stage patients at a dose close to the maximum tolerated dose (MTD) for their therapy, and active experimentation to obtain an accurate estimate of the MTD for use in a subsequent phase II trial. Designs have been proposed to balance these aims, like “continual reassessment” by O`Quigley et al. (1990) and “escalation with overdose control” by Babb et al. (1998). As a benchmark for such designs we consider an optimal sequential design, the computational complexity of which would limit its practical use for even a two parameter dose-response model, but which can be accurately approximated using recent advances in approximate dynamic programming and Monte Carlo simulation. We compare current designs to this near-optimum and discuss what intuition it can provide into the optimal balance between therapy and experimentation.

458 Noninferiority Trial Designs in Medical Devices ● ✪

Biopharmaceutical Section, Biometrics Section, WNAR, ENAR

Wednesday, August 1, 2:00 pm–3:50 pm

Statistics and Clinical, and Regulatory Considerations in the Selection of an Offset for Noninferiority Designs

❖ Philip Lavin, Averion International Corp., 225 Turnpike Road, Southborough, MA 01772, *plavin@averioninc.com*

Key Words: Noninferiority, Regulatory, Offset, Medical Devices, MCID

The choice of an offset (disadvantage to be ruled out) has regulatory, clinical, and statistical motivations and implications. The motivations will be discussed using IDE examples. The implications will be discussed from statistical perspective. There are also hybrid designs where the test device should have a small theoretical advantage, thus dropping sample size. Statisticians also need to be aware of the many computational options for doing binomial calculations since quite wide swings are possible depending on the computational approach used.

Considerations for the Development of a Set of Noninferiority Trials for a Second-Generation Drug Eluting Stent

❖ Joe Bero, Boston Scientific Corporation, One Scimed Place, Maple Grove, MN 55416, *joe.bero@bsci.com*

Key Words: Historical Control, Propensity Adjusted, Non-Inferiority Margin

In the development of drug eluting stent programs, next generation stents are expected to perform as well as previously developed versions of a drug eluting stent. Traditional superiority trials are not useful when comparing similar devices. Noninferiority trials offer a method of comparing devices with similar expected efficacy and safety outcomes while allowing for new devices with desirable characteristics to be brought to market. This will be an examination of the development of a set of trials designed to show non-inferiority of a second generation drug eluting stent compared to a first generation drug eluting stent.

A Noninferiority Trial Utilizing a Closed Hierarchical Testing Procedure

❖ Peggy Pereda, Boston Scientific Corporation, 1 Scimed Place, MS A242, Maple Grove, MN 55311, *peredap@bsci.com*; Eric Bass, Boston Scientific Corporation; Peter Lam, Boston Scientific Corporation

Key Words: Non-Inferiority Trial, Margin, Hierarchical Testing

The goal of the current trial is to assess the optimal revascularization treatment for patients with three-vessel (3VD) or left main (LM) CAD by randomizing them to PCI with DES or CABG. The primary endpoint is 12-month MACCE; the DES group will be assessed for non-inferiority to the CABG group. A closed hierarchical testing procedure will be performed on three populations of interest: all patients, 3VD only, and LM. The primary analysis will be performed on the overall population pooling 3VD and LM. For each population, if the upper one-sided 95% confidence bound for the difference is less than ?, PCI with DES will be considered to be non-inferior to CABG. The order of testing will be: overall, 3VD only and LM. Each test will be performed at a significance level of 5%; the overall family-wise error rate will be controlled at 5% based on the closed hierarchical procedure being used.

Bayesian NonInferiority Trial Design Using Hierarchical Models

❖ Yongyi Yu, Boston Scientific Corporation., 1 Scimed Pl, Maple Grove, MN 55311, *yua@bsci.com*; Frank Zhou, Boston Scientific Corporation; Scott Wehrenberg, Boston Scientific Corporation; Ming-Hui Chen, University of Connecticut

Key Words: Drug-eluting stent, Hierarchical model, Simulation, MCMC, Operating Characteristics

A Bayesian trial is proposed to determine if the new generation drug-eluting stents (DES) is non-inferior to the first generation drug-eluting stents. A hierarchical model is used to incorporate historical information available on first generation DES trials. Study success criteria are based on posterior probability of non-inferiority. Sample sizes, borrowing strategy, and corresponding operating characteristics are simulated and compared to frequentist approach.

Comparison Between Hierarchical Model and Power Prior Approach in a Bayesian Noninferiority Trial Design

❖ Peter Lam, Boston Scientific Corporation, 100 Boston Scientific Way, M1, Biostatistics Dept., Marlborough, MA 01752, *lamp@bsci.com*; Alan Yu, Boston Scientific Corporation; Ming-Hui Chen, University of Connecticut

Key Words: Non-Inferiority, Hierarchical Model, Power Prior, Simulation

Power prior can be a good alternative to commonly used hierarchical approach in Bayesian study design utilizing historical data. Study designs by these two approaches are presented. Differences, similarities, theoretical link and relations between these two approaches are discussed.

459 Shrinkage Theory and Its Applications

IMS, Section on Nonparametric Statistics
Wednesday, August 1, 2:00 pm–4:50 pm

Accounting for Selection Bias in the Estimation of Fold Change in Microarray Experiments

❖ George Wright, National Institutes of Health, NCI, 6130 Executive Blvd Rm 8144, Biometrics Research Branch, Rockville, MD 20852, *wrightge@mail.nih.gov*

Key Words: microarray, shrinkage, empirical bayes

In order to account for the multiple comparisons in microarray experiments it is necessary to choose extremely strict cutoffs for determining statistical significance. As a result those genes selected will tend to include those which have the greatest observed fold changes, through a combination of both chance and true biological difference. Thus the observed fold differences may be much larger than would likely be observed in follow-up experiments that attempt to verify the presented results. The magnitude of this bias depends heavily on the shape tails of the distribution of differential expression. We present several different methods to adjust for this bias and evaluate their performance under a number of simulated and experimental settings.

Objective Bayesian Analysis for Estimating the KL Divergence Between Two Normal Populations with Common Covariance Matrix

❖ Zhonggai Li, Virginia Polytechnic Institute and State University, Dept of Statistics, Hutcheson Hall 503B, Blacksburg, VA 24061-0439, *zhonggai@vt.edu*; Dongchu Sun, University of Missouri-Columbia

Key Words: KL divergence, reference prior, objective prior

In order to compare the two multivariate normal populations with the common covariance matrix, one could consider the Kullback-Leibler (KL) divergence between the two populations, which is proportional to the distance between the two means, weighted by the common precision matrix. In this paper, various objective priors such as the Jeffreys prior, invariance priors and reference priors are studied. Frequentist properties of Bayesian procedures using objective priors are also studied.

Optimal Tests Shrinking Both Means and Variances Applicable to Microarray Data Analysis

❖ Peng Liu, Iowa State University, Department of Statistics, Ames, IA 50011, *pliu@iastate.edu*; Gene J.T. Hwang, Cornell University

Key Words: shrinkage, empirical Bayes, microarary, average power, FDR

As a consequence of large p small n for microarray data, hypothesis tests based on individual genes often result in low average power. There are several proposed tests that attempt to improve power. We studied F_S-test developed with the concept of James-Stein shrinkage to estimate the variance. Under a statistical model providing a theoretical justification of F_S, we derive F_S-test as an empirical Bayes likelihood ratio test. Furthermore, modifying the priors leads to a new test, the F_{MAP} test. An approximation to F_{MAP}, F_{SS}, is derived which was seen to shrink both means and the variances and which has a numerically identical average power as F_{MAP}. Simulation studies show that, the proposed F_{SS} test improves upon many other tests including the classical F test, F_S-test, the test of Wright and Simon, moderated t-test, SAM, Efron's t-test and B-statistics.

Hard and Soft Parameter Shrinkage in Microarray Data Analysis

❖ Gordon Smyth, Walter & Eliza Hall Institute, 1G Royal Parade, Parkville Victoria, 3052 Australia, *smyth@wehi.edu.au*

Key Words: microarray, empirical Bayes, differential expression

This talk will discuss some issues in using parameter shrinkage to overcome the problems of small numbers of arrays when assessing statistical significance in microarray gene expression experiments.

Double Shrinkage Empirical Bayesian Confidence Intervals for Multiple Parameters When the Variances Are Unequal and Unknown

❖ Jing Qiu, University of Missouri-Columbia, 134I Middlebush Hall, Department of Statistics, Columbia, MO 65211, *qiujing@missouri.edu*; Gene J.T. Hwang, Cornell University; Zhigen Zhao, Cornell University

Key Words: Empirical Bayes, Double shrinkage, Multiple confidence intervals, unequal and unknown variances, microarray

In this paper we apply the empirical Bayes technique to construct confidence intervals for a large number p of parameters. The present confidence intervals constructed in the literature assume that the variances σ^2_i's are either known or equal. For the situation when variances are unequal and unknown, the suggestion is typically to replace it by an unbiased estimator S^2_i. However, when p is large, there may be many S^2_i that are extremely small and thus create many false positive results. This problem can be corrected by using shrinkage estimators. Here, we attempt to construct confidence intervals based on empirical Bayes estimators that shrink both the means and variances. Analytical and numerical studies and application to a real data show that compared to t intervals, our intervals have higher coverage probabilities while yielding shorter lengths.

460 Monte Carlo Methods and Their Applications

IMS, General Methodology

Wednesday, August 1, 2:00 pm–3:50 pm

Improving SAMC Using Smoothing Methods: Theory and Applications

❖ Faming Liang, Texas A&M University, Department of Statistics, College Station, TX 77843, *fliang@stat.tamu.edu*

Key Words: Model selection, Stochastic Approximation, smoothing, Monte Carlo

Stochastic approximation Monte Carlo (SAMC) has recently been proposed by Liang et al (2006) as a general simulation and optimization algorithm. In this talk, we propose to improve its convergence using smoothing methods and discuss the application of the new algorithm to Bayesian model selection problems. Our numerical results show that the improvement is significant. The new algorithm represents a general form of the stochastic approximation Markov chain Monte Carlo algorithm. It allows multiple samples to be generated at each iteration, and a bias term to be included in the parameter updating step. A rigorous proof for the convergence of the general algorithm is established under verifiable conditions. This paper also provides a framework on how to incorporate nonparametric techniques into Monte Carlo methods to improve Bayesian computation.

Sampling Issues for Largish Tables

❖ Ian Dinwoodie, Duke University, ISDS Box 90281, Durham, NC 27708, *ihd@stat.duke.edu*; Yuguo Chen, University of Illinois at Urbana-Champaign

Sampling multi-way contingency tables with margin constraints can be difficult for certain types of constraints on large tables. We will discuss some strategies to improve the quality of the sample on specific examples where the Markov basis is hard to compute.

Ensemble Sampling

❖ Hedibert Lopes, The University of Chicago Graduate School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, *hlopes@chicagogsb.edu*; Nicholas Polson, The University of Chicago Graduate School of Business

Key Words: State-Space Models, Stochastic Volatility, Multivariate, Markov Chain Monte Carlo, Kalman filter, Sequential Monte Carlo

In this project we provide a robust ensemble filtering and learning algorithm. Ensemble filtering methods are based on an update rule for particle propagation rather than resampling methods. Hence, the methods scale to high dimensions and avoid particle degeneracy. In the case of pure state filtering we compare our filter in a univariate stochastic volatility model with an the existing methodologies. We also provide a multivariate application to factor multivariate stochastic volatility.

Interweaving Data Augmentation Schemes for Faster MCMC

❖ Yaming Yu, University of California, Irvine, 346A ICS Bldg One, Department of Statistics, Irvine, CA 92697-1250, *yamingy@ics.uci.edu*

Key Words: Data Augmentation, Missing Data, MCMC, GLMM

Re-parameterizations, or variable transformations, are known to be a key to efficient implementation of Gibbs sampling or general MCMC algorithms. Rather than focusing on a single parameterization, we advocate the method of interweaving two or more data augmentation schemes and demonstrate the benefit of interweaving two special schemes, sufficient augmentation and ancillary augmentation, in applied settings. This is a continuation of Yu and Meng (2007).

The Evolutionary Forest

❖ Leman Scotland, Duke University, 1800 Shelton Ave, Durham, NC 27707, *scotland@stat.duke.edu*

Key Words: Bayesian, MCMC, Phylogeny, Population Biology

We describe a Markov chain Monte Carlo method for approximating the joint posterior distribution of parameters for evolutionary and other complex processes. Proposal distributions on complex structures such as phylogenies are essential for MCMC sampling methods. However, such proposal distributions are difficult to construct so that their probability distribution match that of the true target distribution, in turn hampering the efficiency of the overall MCMC scheme. We'll describe a data augmentation scheme that converges rapidly to the population parameters of interest, while utilizing a simple independent proposal distribution on individual trees. In this approach, we rely on an ensemble of histories (a forest of genealogical trees) rather than a single history. This enables the exploration of the augmented tree space (forest space) to proceed quickly and converge rapidly.

461 The Current Trend of Bayesian Model Comparison and Variable Selection ● ✪

Section on Bayesian Statistical Science, ENAR

Wednesday, August 1, 2:00 pm–3:50 pm

Objective Bayesian Model Selection for a Linear Mixed Model

❖ Dongchu Sun, University of Missouri-Columbia, 134C Middlebush Hall, Department of Statistics, Columbia, MO 65211, *sund@missouri.edu*

Key Words: Bayes factor, Objective priors, Consistency, Random effect models

It is widely accepted that the Bayes factor is a key tool in model selection. Nevertheless, it is an important, difficult question which priors should be used. We first consider this problem in the context of the one-way random effects model. Arguments based on concepts like orthogonality, matching predictive, and invariance are used to justify a specific form of the priors, where the prior for the new parameter has to be determined. Two different proposals for this proper prior have been derived: the intrinsic priors and the divergence based priors, a recently proposed methodology. It is shown that the divergence based priors produce consistent Bayes factors. The methods are illustrated on examples and compared with other proposals. The method for a general linear mixed model is also explored. (Jointly with Gonzalo Garcia-Donato, University of Valencia, Spain).

An Extended BIC for Model Selection

❖ Surajit Ray, Boston University, 111 Cummington St, Boston, MA 02215, *sray@bu.edu*; James Berger, Duke University; Susie Bayarri, University of Valencia; Woncheol Jang, University of Georgia; Luis R. Pericchi, University of Puerto Rico, San Juan

Key Words: Bayes factor, Cauchy Priors, Consistency, Model Selection, Effective sample size, Fisher Information

We present a new approach to Bayes factors based on Laplace expansions (as BIC) which we call EBIC (Extended BIC). In our approach, we do not include the prior in the Laplace expansion, but choose it appropriately so that it produces close-form expressions for the resulting EBIC. We explore both joint priors and independent priors for the component parameters. To help choose the scale of the prior, we use a novel definition of effective sample size which allows for different effective sample sizes for the parameters. The new EBIC avoids many of the difficulties commonly associated with BIC, and can often be shown to be consistent. We also produce a modified EBIC which is more favorable to complex models while still retaining consistency. A comparative study of a range of model selection tools including the EBIC will be presented during the talk.

Random Effects Selection with Nonparametric Settings

❖ Bo Cai, University of South Carolina, 800 Sumter Street, suite 205, Columbia, SC 29208, *bocai@gwm.sc.edu*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Dirichlet process, Latent variables, Nonparametric Bayes, Stochastic search, Subset selection

In analyzing longitudinal or clustered data with a mixed effects model, violations of normality can potentially impact subset selection, inferences on the heterogeneity structure, and the accuracy of predictions. Although there is often interest in identifying the subset of predictors having random effects, random effects selection can be challenging, particularly when outcome distributions are non-normal. This article focuses on Bayesian methods for subset selection in nonparametric random effects models in which one is uncertain about the predictors to be included and the distribution of their random effects. By using carefully-chosen mixture priors for coefficients in the base distributions of the component Dirichlet Process, we allow fixed and random effects to be effectively dropped out of the model. An approximation is used for intractable integrals for non-normal outcomes.

Objective Bayesian Variable Selection for Logistic Regression Models with Jeffreys' Prior

❖ Ming-Hui Chen, University of Connecticut, Department of Statistics, 215 Glenbrook Road, U-4120, Storrs, CT 06269-4120, *mhchen@stat.uconn.edu*; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill; Sung Duk Kim, University of Connecticut

Key Words: BIC, Importance sampling, Normalizing constant, Posterior probability, Tail behavior, Variable selection

We study several theoretical properties of Jeffreys' prior for logistic regression models with a focus on its applications to variable selection problems. We show that Jeffreys' prior is symmetric and unimodal about 0 and always has lighter tails than a t distribution and heavier tails than a normal distribution. We also develop an efficient importance sampling algorithm for calculating the prior and posterior normalizing constants based on Jeffreys' prior. Moreover, we show that the prior and posterior normalizing constants under Jeffreys' prior are scale invariant in the covariates. A closed form for Jeffreys' prior is obtained for saturated logistic regression models with binary covariates. Detailed simulation studies are presented to demonstrate its properties and performance and a real dataset is also analyzed to further illustrate the proposed methodology.

462 Reducing Disclosure Risk Through Statistical Methods and Administrative Procedures ✪

Section on Government Statistics, Section on Health Policy Statistics, Social Statistics Section

Wednesday, August 1, 2:00 pm–3:50 pm

Comparative Evaluation of Four Sensitive Tabular Data Protection Methods Using a Real-Life Table Structure of Complex Hierarchies and Links

❖ Ramesh Dandekar, Energy Information Administration, 8922 Applecross Lane, Springfield, VA 22153-1201, *ramesh.dandekar@gmail.com*

Key Words: LINKED TABLES, HIERARCHICAL TABLES, SUPPRESSION, NOISE ADDITION, CONTROLLED TABULAR ADJUSTMENTS, network flow models

The practitioners of tabular data protection methods in federal statistical agencies have some familiarity with commonly used table structures. However, they require some guidance on how to evaluate appropriateness of various sensitive tabular data methods when applied to their own table structure. With that in mind, we use a real life "typical" table structure of moderate hierarchical and linked complexity and populate it with synthetic micro data to evaluate relative performance of four different tabular data protection methods. The methods selected for the evaluation are: 1) classical cell suppression 2) lp-based CTA (Dandekar 2001), 3) USBC's network flow-based cell suppression and 4) USBC's micro data level noise addition method. The outcome from the comparative evaluation is available from http://mysite.verizon.net/vze7w8vk.

Implementing a Portfolio Approach to the Protection of Confidential Data

❖ Tim Mulcahy, NORC at the University of Chicago, 1350 Connecticut Ave NW, Washington, DC 20036, *mulcahy-tim@norc.org*; Julia Lane, NORC at the University of Chicago

Key Words: confidentiality, data dissemination

US federal agencies and departments disseminate statistical data to external researchers for a number of reasons. These include allowing external researchers to conduct analyses in areas of interest to the agencies, to identify data quality issues, or to find new and innovative research and educational uses for existing datasets without further burdening respondents or increasing agency costs. The view expressed by a recent panel of the National Commission on Health Statistics about the best way to promote high quality academic research is shared by many: panel members explicitly stated that researchers needed to access and view original microdata directly in their offices, conduct their analyses, share their results, and engage in discourse about different aspects of the data, the analysis, and the interpretation of the results. Smith (1991) makes a similarly compelling case.

Matching NCES Data to External Databases To Assess Disclosure Risk

Elizabeth Steiner, Carnegie Mellon University; ❖ J. Neil Russell, National Center for Education Statistics, 1990 K. St., NW, Room 9050, Washington, DC 20006, *neil.russell@ed.gov*; Marilyn Seastrom, National Center for Education Statistics

Key Words: Confidentiality, Matching, External Databases, Disclosure Risk

By law, the National Center for Education Statistics (NCES) is required to protect respondent identifying information. Since the 1990s, the NCES Disclosure Review Board (DRB) has required that survey programs that release public-use microdata files (PUMFs) match to external databases as part of a disclosure risk analysis. Since there are few published accounts of Federal agencies using matching as a method of assessing disclosure risk, this paper contributes to the literature in two ways: 1) this paper will chronicle the NCES history of matching as a disclosure risk assessment method, and 2) it will present disclosure risks discovered by matching to external databases. The paper provides a brief history of NCES matching studies, describes the methods used to conduct this meta-analysis, and presents findings in the context of assessing disclosure risk.

Analytically Valid Microdata Files and Re-identification

❖ William Winkler, U.S. Census Bureau, 7705 Heritage Drive, Annandale, VA 22003, *william_e_winkler@msn.com*

Key Words: record linkage, masking, synthetic microdata

With the exception of synthetic data (e.g., Reiter 2002, 2005) and a few other methods (Kim 1986, Dandekar, Cohen, and Kirkendal 2002), masking methods and resultant public-use files are seldom justified in terms of valid analytic properties. If a file has valid analytic properties, then the analytic characteristics can be used as a starting point for re-identification using analytic methods only (Lambert 1993, Fienberg 1997). In this paper, we describe a general method for building a synthetic data file having valid analytic properties. While re-identification in the synthetic files is primarily based on analytic properties, it is possible to speed up and increase re-identification using a combination of analytic and computerized record linkage procedures.

Measuring Disclosure Risk and an Examination of the Possibilities of Using Synthetic Data in the Individual Income Tax Return Public Use File

❖ Michael Weber, Internal Revenue Service, Statistics of Income, 4545 Windsor Lane, Bethesda, MD 20814, *michael.e.weber@irs.gov*; John Czajka, Mathematica Policy Research, Inc.; Sonya Vartivarian, Mathematica Policy Research, Inc.

Key Words: Synthetic data, Public Use, Income Tax Returns

The Statistics of Income Division (SOI) currently measures disclosure risk through a distance based technique that compares the Public Use File against the population of all tax returns and uses top-coding, subsampling and multivariate microaggregation as disclosure avoidance techniques. SOI is interested in exploring the use of other techniques that prevent disclosure while providing less data distortion. Synthetic or simulated data may be such a technique. But while synthetic data may be the ultimate in disclosure protection, creating a synthetic dataset that preserves the key characteristics of the source data presents a significant challenge. An additional constraint in creating synthetic data for the SOI PUF is found in maintaining the accounting relationships among numerous income, deduction, and tax items that appear on a tax return.

463 Nonparametric and Semiparametric Methods in Brain Imaging Studies

Section on Nonparametric Statistics, International Indian Statistical Association

Wednesday, August 1, 2:00 pm–3:50 pm

Multisubject Independent Component Analysis Using a Maximum Likelihood Approach

❖ Ying Guo, Emory University, 1518 Clifton RD NE, Department of Biostatistics, Atlanta, GA 30322, *yguo2@sph.emory.edu*

Key Words: independent component analysis, functional magnetic resonance imaging, Multi-subject ICA, maximum likelihood approach, likelihood ratio tests, test of group differences

Independent component analysis (ICA) has become increasingly popular for analyzing functional magnetic resonance imaging (fMRI) data. The extension of ICA for group inferences is an active research topic. Existing group ICA models assume different underlying structures of group spatio-temporal processes. There is currently no method for assessing the appropriateness of the assumed structure in a data set. Another challenge in group ICA analysis is how to test differences between subject groups. I propose a maximum likelihood approach which offers a unified framework for estimating and comparing group ICA models with different underlying structures. A class of likelihood ratio tests are derived to assess the goodness-of-fit of a group ICA model and to test the homogeneity in ICA between two groups. Simulation studies and application to a fMRI data example would be discussed.

Detecting Activations in fMRI Experiments with Maximum Cross-Correlation Statistics

❖ Kinfemichael Gedif, Southern Methodist University, 6236 Ridgecrest Rd Apt 2424, Dallas, TX 75231, *kgedif@smu.edu*; William R. Schucany, Southern Methodist University ; Richard Gunst, Southern Methodist University; Wayne Woodward, Southern Methodist University; Jeffrey Spence, The University of Texas Southwestern Medical Center at Dallas;

Patrick Carmack, The University of Texas Southwestern Medical Center at Dallas; Qihua Lin, The University of Texas Southwestern Medical Center at Dallas

Key Words: fMRI, cross-correlation, HRF, kernel-smoothing

Cross-correlation between an anticipated hemodynamic response function (HRF) and the observed voxel's time course is one method of identifying activated voxels due to the known stimulus sequence in an fMRI experiment. In fact, such a method is an optimal detector if the underlining noise process is Gaussian and the true signal of interest is known. In practice however, the true signal of interest is not known and analysis based on fitting HRF models strongly depends on the adequacy of the fitted model. We propose an approach that does not require fitting an HRF to the voxel time series. The maximum cross-correlation between the kernel-smoothed ideal stimulus sequence and shifted (lagged) values of the observed response is the proposed test statistic. Improved efficiency is demonstrated in simulations involving realistic drift and noise models.

Thresholding Methods Applied on fMRI Data

❖ Martina Pavlicova, Columbia University, 722 W 168th str 6th fl, New York, NY 10032, *mp2370@columbia.edu*

Voxel-wise analyses are common methods for processing fMRI data. Because of the large number of hypotheses involved, finding an activation threshold is a multiple-comparisons problem. Bonferroni comparisons and cluster-based thresholding are frequently used for this purpose. We propose using a modification of a procedure that controls the proportion of false positives among the voxels declared to be activated, namely the False Discovery Rate (FDR).

A Bootstrap Resampling Approach: Tests of Significance on Brain Imaging Data

❖ Chung Chang, Columbia University, 220 East 70th Street, Apt 9C, New York City, NY 10021, *cc2240@columbia.edu*; R. Todd Ogden, Columbia University

Key Words: multiple comparisons, spatial correlation structure, SPM, permutation test, bootstrap, family-wise error rate

In neuroimaging studies it is of interest to test for changes in imaging data among subjects in different groups. Testing hypotheses voxel by voxel results in a multiple comparisons problem for which solutions should take into account the spatial correlation structure inherent in the imaging. Statistical Parametric Mapping (SPM) and the permutation test have become popular in this setting but they rely on parametric and exchangeability assumptions, respectively, which are not always satisfied in practice. We propose a bootstrap approach (L1) that is free of the parametric assumptions made by SPM and also are more flexible than the permutation test. We compare the performance of the L1 method with that of SPM, the permutation test, and another bootstrap approach (L2). For the L2 method, we present sufficient conditions that ensure asymptotic control of the family-wise error rate.

Bayesian Hierarchical Spatiotemporal Models for fMRI Data

❖ Qihua Lin, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, PO Box 8896, Dallas, TX 75390-8896, *catherine.lin@utsouthwestern.edu*; Richard Gunst, Southern Methodist University; Jeffrey Spence, The University of Texas Southwestern Medical Center at Dallas; Patrick Carmack, The University of Texas Southwestern Medical Center at Dallas; William R. Schucany, Southern Methodist University

Key Words: fMRI, Spatiotemporal, Bayesian

A rich spatial and temporal correlation structure is well-known to be inherent in fMRI data and the signal-to-noise ratio is generally low. A class of spatiotemporal models is introduced to model such data. Both low-frequency drift and hemodynamic response functions with spatially varying parameters are modeled to explain large-scale variations. Spatially and temporally correlated small-scale variations are modeled by autoregressive moving average processes. Parameters in this class of models are estimated in a Bayesian hierarchical setting, which provides a convenient way to account for spatial correlations. This method offers a strong alternative to SPM for analyzing fMRI data.

464 Dimension Reduction that Informs Science ● ✪

Section on Physical and Engineering Sciences

Wednesday, August 1, 2:00 pm–3:50 pm

Model Selection and Estimation for Semiparametric Stochastic Mixed Models for Longitudinal Data

❖ Xiao Ni, North Carolina State University, Department of Statistics, Campus Box 8203 NCSU, Raleigh, NC 27695, *xni@stat.ncsu.edu*; Daowen Zhang, North Carolina State University; Hao Zhang, North Carolina State University

Key Words: correlated data, Gaussian stochastic process, variable selection, smoothly clipped absolute deviation, smoothing splines

We propose a double-penalized likelihood approach for simultaneous model selection and estimation in semiparametric stochastic mixed models for longitudinal data. Two penalties are imposed on the ordinary likelihood: the roughness penalty for the nonparametric baseline function and the shrinkage penalty for the linear coefficients. We present an algorithm for maximizing the double-penalized likelihood and solving for various model components and unknown parameters within a modified linear mixed model framework. We propose and compare frequentist and Bayesian inference for model parameters. Simulation results are given to show effectiveness of our method. A real data example is also given to illustrate the use of our method.

Shrinkage Inverse Regression for Model-Free Variable Selection

❖ Lexin Li, North Carolina State University, Department of Statistics, Raleigh, NC 27695, *li@stat.ncsu.edu*; Howard D. Bondell, North Carolina State University

Key Words: Inverse regression estimation, Nonnegative garrote, Sufficient dimension reduction, Variable selection

The family of inverse regression estimators recently proposed by Cook and Ni (2005) have proven effective in dimension reduction by transforming the high-dimensional predictor vector to its low-dimensional projections. In this talk, we propose a general shrinkage estimation strategy for the entire inverse regression estimation family, which is capable of simultaneous dimension reduction and variable selection. We demonstrate that the new estimators achieve consistency in variable selection without requiring any traditional model, meanwhile retaining the root-n estimation consistency of the dimension reduction basis. We also show the effectiveness of the new estimators through both simulation and real data analysis.

Gene Expression Data Analysis Using the Gene Ontology

❖Jiajun Liu, North Carolina State University, 3009 Apt C Dorner Circle, Raleigh, NC 27606, *jliu6@stat.ncsu.edu*; Jacqueline M. Hughes-Oliver, North Carolina State University; J. Alan Menius, GlaxoSmithKline

Key Words: Microarray, Gene Ontology, PLS, Domain aggregation

New technologies for biological systems give scientists the ability to record thousands of measurements for each biomolecule including genes, proteins and metabolites. Domain enhanced analysis (DEA) uses the Gene Ontology to guide analysis of such data and increase interpretability. DEA uses a "top-down" approach to perform domain aggregation by first combining gene expressions related to each GO term using the Partial Least Squares (PLS) procedure. Then the first scores from the PLS procedure are used to test for differentially expressed patterns using the t test. We find the general t test inadequate for adjusting for the number of genes within each GO term. New tests are proposed by finding the true null distribution of each PLS score adjusted for the size of the GO term. We also discuss the impact of using different two-class classification response variables, namely 0/1 or -1/1.

Recent Directions in Predictor Selection

❖Jessica Kraker, University of Wisconsin-Eau Claire, 216 Helen Street, Roberts, WI 54023, *krakerjj@uwec.edu*; Douglas M. Hawkins, The University of Minnesota

Key Words: predictor selection, penalized regression, chemometrics

Beginning with the closed-form ridge regression model (quadratic loss and penalty) and advancing to more computationally-intensive methods (such as the lasso and elastic net), the possibilities for penalized regression have progressed dramatically in recent years. In the context of chemometrics, we analyze a prediction problem calling for the concurrent selection of predictors with fitting of the regression model. Model selection from among several penalized regression models (with different loss and penalty functions) requires the further assessment of the model utility. Models are included with consideration for the type of predictors implemented by the researcher and for the type of loss function desired. Programming is implemented in the R environment to obtain and to assess the fitted models.

Dimension Reduction in Systems Biology

❖Lei Zhu, GlaxoSmithKline, 5 Moore Drive, MAI T205, Research Triangle Park, NC 27709, *lei.a.zhu@gsk.com*; Kwan Lee, GlaxoSmithKline; Amit Bhattacharyya, GlaxoSmithKline; Edit Kurali, GlaxoSmithKline; Amber Anderson, GlaxoSmithKline

Key Words: Multivariate Analysis, Omics Platform, Selection Bias, Supervised Analysis, Unsupervised Analysis, Univariate Analysis

Pharmaceutical industry is facing critical challenges in translational efficacy failure. Systems biology emerged as a paradigm shift from target-focused drug discovery to a systems approach. The automated acquisition of large amounts of omics data creates exploratory and interpretative analysis challenges in our attempt to unravel associations among blood chemistry, transcripts, proteins, metabolites, and lipids under the drug or disease perturbations. This paper discusses the various analysis approaches for dimension reduction on systems biology data to identify drug and/or disease biomarkers for better understanding of drug efficacy. The analysis approaches fall into four categories: univariate unsupervised or supervised, multivariate unsupervised or supervised. Application examples of these approaches at different stages of data analysis will be demonstrated.

465 Innovative Ideas for Teaching Statistics

Section on Statistical Education, ENAR, Section on Teaching Statistics in the Health Sciences
Wednesday, August 1, 2:00 pm–3:50 pm

Teaching Introductory Statistics with a Blended Method Approach

❖Leonard Gaines, Empire State Development, 30 S Pearl St, Albany, NY 12245, *lgaines@empire.state.ny.us*

Key Words: Distance Learning, Adult students

SUNY-Empire State College's FORUM program is designed for adults working to earn their Associate and Bachelor degrees while working a full-time job. Students in this program attend three weekend residencies each term. During this time, their classes meet for three hours, providing nine hours of classroom contact with their classmates and faculty member each term. In between these residencies, they maintain contact with their instructor in a variety of ways. This presentation describes how the residencies are combined with an Internet-based approach in an introductory statistics course. Observations about the effectiveness of different activities in this blended-learning model in comparison to similar learning activities in a purely distance learning Internet-based version of the course will be made.

Teaching Introductory Statistics with Constraints and High Expectations

❖William Seaver, University of Tennessee, Dept of Statistics and OMS, College of Business Administration, Knoxville, TN 37966-0532, *wseaver@utk.edu*

Key Words: Teaching, Introductory statistics, Reducing variation, Recruiting for statistics

About 2500 students a year take our introductory statistics course, and that amount is increasing 5% a year. About 40% are business majors, 20% are business minors, and the rest are from disciplines all across campus. About three years ago, some disturbing issues were occurring, too much variation. In addition, instructor changes from faculty only to some faculty and master's students and now to Ph.D. students with an occasional faculty member needed to be addressed. The urgency of the situation demanded some immediate change along with some long-term changes. Our department, our college, and the university have been excited about the progress we are making in light of the constraints. As a result, the students are learning statistics better; and there is a waterfall affect on other courses, majors, and minors. This paper addresses this journey and solutions!

To Service Learn (SL) or Not To Service Learn?

❖Amy Phelps, Duquesne University, Rockwell Hall, 600 Forbes Avenue, Pittsburgh, PA 15282, *phelpsa@duq.edu*

Key Words: comparative study, service learning, statistics projects, authentic assessment

This study gave students the choice of participating in a SL project or doing a traditional project assignment. Both options were designed to support the GAISE guidelines. Did the SL activity enhance the learning experience? When the projects were completed, students answered a survey designed to reflect on the experience. Both groups responded equally 'agree or strongly agree' to the Likert scale questions: 96.15% reinforced learning objectives, 98.08% applied to real world, 84.62% positive experience. Responses to the open ended questions: more students in the SL group (p = 0.019) wrote about the benefits of dealing with real world data, more SL students felt their work benefited others (65% felt their statistical expertise was valuable) and more (p=0.005) SL students felt that the experience will help them in future classes.

Engaging Students in Elementary Statistics

❖ Greg Knofczynski, Armstrong Atlantic State University, 11935 Abercorn Ext, Savannah, GA 31419, *knofczgr@mail.armstrong.edu*; Paul Hadavas, Armstrong Atlantic State University; Lorrie Hoffman, Armstrong Atlantic State University

Key Words: education, elemenatary statistics, student perspective

In hopes of reducing the percent of students receiving non-passing grades in elementary statistics classes, interactive hands-on projects were implemented in the classes. These projects gave students the opportunity to use personal data, discuss statistical concepts with each other, and strengthen the students' understanding, mastery and appreciation of the material covered in an elementary statistics class. Although no decrease in the percent of non-passing grades was observed and no noticeable shift in the overall grades was observed, there was a statistically significant increase in the students' perspectives of the teacher's ability to teach the course.

Enhancing the Teaching of Statistics with Technology Using the Statistics Online Computational Resource (SOCR)

❖ Nicolas Christou, University of California, Los Angeles, 8125 Math Sciences Bldg, Los Angeles, CA 90095, *nchristo@stat.ucla.edu*; Ivo D. Dinov, University of California, Los Angeles; Juana Sanchez, University of California, Los Angeles

Key Words: education research, simulation applets, teaching with technology, java applets, online course materials, probability and statistics

The NSF-funded Statistics Online Computational Resource (www.SOCR.ucla.edu) provides a number of educational materials and interactive tools for enhancing instruction in various undergraduate and graduate courses in probability and statistics. SOCR includes class notes, practice activities, statistical calculators, interactive graphical user interfaces, computational and simulation applets, tools for data analysis and visualization. Based on the promising results from our pilot study in 2005–2006, where we saw a consistent trend of improvement in classes using SOCR compared to those not using it, in terms of quantitative examination measures, we present in this talk the results of a new study contacted in 2006–2007. The goal of this new study is to attempt to associate the effects of SOCR with students' individual learning style based on the Felder-Silverman-Solomon index.

466 Computationally Intensive Methods in AIDS Research II ● ✪

WNAR, ENAR, Section on Health Policy Statistics, Section on Nonparametric Statistics

Wednesday, August 1, 2:00 pm–3:50 pm

Nonparametric Tests with Application to RNA Viral Load Decline

❖ Susanne May, University of California, San Diego, Division of Biostat and Bioinf, 9500 Gilman Dr MC0717, La Jolla, CA 92093, *smay@ucsd.edu*; Victor De Gruttola, Harvard University

Key Words: Censored biomarker measurements, Longitudinal Data, Resampling, U-Statistic

We propose new tests for two-group comparisons of repeated measures that might be obtained at arbitrary time points and differ over individuals. The tests do not make any assumptions regarding the distribution of the repeated measures except that one of them assumes that the repeated measures can be grouped into distinct periods of observations such that the covariance between scores only depends on the periods the observations belong to and that the covariance matrices are the same in the two groups. The tests remain valid even if the probability that a response is observed depends on the level of response if missing data mechanisms are the same in both groups. Inference can be based on resampling. We use the tests to assess differences in viral load decline for drug resistant and drug sensitive human immunodeficiency virus (HIV)-1 infected patients.

Non- and Semiparametric Subset and Model Selection for Multiple Categorical Predictors and Several Outcomes

❖ A. G. DiRienzo, Harvard School of Public Health, 655 Huntington Avenue, Department of Biostatistics, Boston, MA 02115, *dirienzo@hsph.harvard.edu*

Key Words: Genomics, Genotype sequence, Multiple hypothesis testing, Multiway combinations, Simultaneous inference, Sparse data

Consider selection of a semiparametric model for response given multiple categorical variables. It is wished that the model is easily interpretable and potentially includes higher-order predictor variable/level combinations. Presented is a two-stage approach for analysis. The first stage is a nonparametric subset selection technique that estimates the set of predictor combinations that have a nonredundant marginal association with response. The subset selection complexity parameters correspond to two error rates (e.g. (generalized) familywise error rate, proportion of false positives) and thus are easily interpretable and provide control over the number or proportion of false positives. The second stage constructs a simultaneous confidence band for the prediction accuracy of each member in a set of varying-sized candidate models. Application to a recent AIDS clinical trial is provided.

Survey of Recombination in the HIV Genome: Hot Spots and Cold Spots

❖ Karin Dorman, Iowa State University, Department of Statistics, Ames, IA 50011, *kdorman@iastate.edu*; Misha L. Rajaram, Iowa State University; Marc A. Suchard, University of California, Los Angeles; Vladimir N. Minin, University of California, Los Angeles

Key Words: Bayesian phylogenetics, Markov chain Monte Carlo, change point model

Coinfection of a single cell with two or more HIV strains may produce recombinant viruses upon template switching by the replication machinery. We used a Bayesian multiple change point model to infer inter-subtype recombination occurring during the worldwide epidemic of HIV as recorded in Genbank. We examined 4074 sequences identified as recombinant in a preliminary screen and confirmed 2360 with strong evidence of recombination. These recombinants conservatively contained 1701 unique crossover points, but there is evidence that some co-located crossover points probably descended from distinct recombination events. Per-site recombination rate estimates confirmed the presence of a known hotspot in the envelope gene and a previously hypothesized hotspot in the gag gene. In contrast to previous findings, recombination rate was positively correlated with fast-evolving genomic regions.

Numeric EM for Linear and Nonlinear Mixed Effects Models with Censored Responses

❖Lin Liu, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0717, *linliu@ucsd.edu*; Florin Vaida, University of California, San Diego

Key Words: detection limit, nonlinear models, truncated normal

Laboratory-based data are often censored due to limitations of measuring technology. In AIDS research, HIV viral load is subject to both upper and lower detection limits. The usual EM estimation procedure for mixed effects model needs to be modified to deal with such data issue. We propose a numeric implementation of the EM algorithm for linear/nonlinear mixed effects models with censored response. It is shown that the E-step reduces to computing the first two moments of certain truncated multinormal distributions. We use a numerical implementation which greatly improves the speed and precision in computing the MLE compared to the Monte Carlo EM algorithm. The likelihood function can be calculated as a bi-product of E-step and is used for monitoring convergence. We apply this work in the analysis of the longitudinal viral load of a cohort of subjects with primary HIV infection.

467 Bayesian Approaches to Clinical Trials ●

Section on Bayesian Statistical Science, Biopharmaceutical Section
Wednesday, August 1, 2:00 pm–3:50 pm

Monitoring Late Onset Toxicities in Phase I Trials Using Predicted Risks

❖Benjamin N. Bekele, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Box 447, Houston, TX 77030, *bbekele@mdanderson.org*

Key Words: Adaptive Design, Bayesian Inference, Discrete Time Model, Isotonic Regression, Predictive Probability

Late onset toxicities are a serious concern in phase I trials in oncology. Since most toxicities occur soon after the start of therapy, conventional dose-finding methods rely on a binary indicator of toxicity occurring within a relatively short initial time period. If an agent causes late onset toxicities, such methods may allow an undesirably large number of patients to be treated at toxic doses before any toxicities are observed. We use a Bayesian dose-finding method based on toxicity time data. We use a set of rules, based on predictive probabilities, which may suspend accrual if the risk of toxicity is too high. If additional follow up data later reduce the predicted risk of toxicity to an acceptable level, accrual is re-started, and the process is repeated several times. An extensive simulation study shows that our proposed method provides a measure of safety over other methods.

Decision Analysis in Support of Clinical Trial Protocol Decisions

❖Joseph Kahn, Novartis Pharmaceuticals, 320 South Street 17F, Morristown, NJ 07960, *joseph.kahn@novartis.com*

Key Words: Decision Analysis, Adaptive design, Bayesian, Clinical, Trial, Protocol

Presentation will illustrate via a case study for a phase IIa/b adaptive program design how modelers can use decision analysis and Bayesian modeling to more productively assist clinical teams with protocol decisions. Topics include framing and developing a manageable set of compelling alternatives, assuring information is relevant and has credence, salient value measures, analytic methods useful for implementation, and achieving buy-in from both clinical teams and colleagues in statistics.

Could Bayesians and Frequentists Agree When Designing Clinical Trials?

❖Lurdes Inoue, University of Washington, F600 Health Sciences Building, Seattle, WA 98195, *linoue@u.washington.edu*

Key Words: Bayesian design, clinical trial, sample size determination, power, utility function

In this talk we will discuss the design of a clinical trial considering one of its most critical aspects - sample size determination. We will describe a framework where we will examine sample size calculation under frequentist and Bayesian approaches. Despite the philosophical differences, we will show correspondences between Bayesian and frequentist designs.

Bayesian Optimal Design in Phase II Studies

❖Gary Rosner, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 447, Houston, TX 77030, *glrosner@mdanderson.org*; Peter Mueller, The University of Texas M.D. Anderson Cancer Center; Meichun Ding, Amgen Inc.

Key Words: Optimal design, Phase II clinical study, Sequential design

In anticancer drug development, phase II studies screen out new therapies if they show activity. The treatments that pass the phase II test will undergo further evaluation, such as in a phase III clinical trial. In general, these phase II designs consider each treatment in isolation, with historical information entering the design through specification of null and alternative hypotheses. We propose a systematic decision-making approach to the phase II screening process. We discuss optimal and approximately optimal Bayesian sequential designs for phase II studies screening for active drugs. Computer simulations show that the methodology leads to high probability of discarding treatments with low success rates and moving treatments with high success rates to phase III trial.

Incorporating Patient's Characteristics in Cancer Phase I Clinical Trials Using Escalation with Overdose Control

❖Mourad Tighiouart, Emory University, 1365-B Clifton Rd, Room B 4111, Atlanta, GA 30322, *mourad.tighiouart@emoryhealthcare.org*; André Rogatko, Emory University; Zhiheng Xu, Emory University

Key Words: EWOC, MTD, Cancer phase I clinical trials, Bayesian design

We describe a design for cancer phase I clinical trials that takes into account patients heterogeneity thought to be related to treatment susceptibility. The goal is to estimate the maximum tolerated dose given a set of patient specific covariate values. The design is Bayesian adaptive based on escalation with overdose control paradigm (EWOC). In the case of a binary covariate, we will assess the performance of this method by comparing the following designs via extensive simulations: (1) Design using a covariate,(2) Design ignoring the covariate, and (3) Design using separate trials. A similar comparison is made for a continuous covariate. Since we do not know whether or not the MTD depends on a given patient characteristic in general, we investigate a procedure for determining the minimum sample size needed in order to decide whether or not to include the covariate during the trial.

468 Recent Advance in Analysis of Incomplete Data ● ✪

Biometrics Section

Wednesday, August 1, 2:00 pm–3:50 pm

Examining Benefits and Drawbacks of Multiple Imputation for Disease Rate and Demographic Data

❖ Irene Helenowski, Northwestern University, 680 N Lake Shore Drive, Suite 1102, Chicago, IL 60646, *i-helenowski@northwestern.edu*; Hakan Demirtas, University of Illinois at Chicago; Chiang-Ching Huang, Northwestern University

Key Words: multiple imputation, spatial statistics, parameter estimation

Missing data could be prevalent in spatial problems, as in geographic studies and examining spatial reproducibility of biomarkers in a particular organ. Multiple imputation methods based on the sample distribution of observed data or bootstrapping predicted values from a regression model have been used to circumvent this problem. In this work, plausibility of the imputation approach is examined via measures of spatial discrepancies calculated through multiply imputed datasets. We examine the reliability of the multiple imputation methods described above to study the impact of demographic factors on cancer and cardiovascular disease rate in different geographic areas by examining estimated regression parameters with commonly accepted accuracy and precision measures that involve coverage rate and average width for different numbers of imputed datasets.

Imputation Based on Propensity Score: A Comparative Study of the Penalized Spline Propensity Prediction Method with Alternative Double Robust Estimators

❖ Guangyu Zhang, University of Michigan, 3566 Green Brier Blvd Apt 412C, Ann Arbor, MI 48105, *guangyuz@umich.edu*; Roderick J. Little, University of Michigan

Key Words: missing data, imputation, PSPP, propensity score

Little and An (2004) proposed the penalized spline propensity prediction method of imputation (PSPP). The predicted unconditional mean of the missing variable has a double robustness (DR) property under mis-specification of the imputation model. The DR property can also be achieved by modeling the relationship parametrically. One method is to include the inverse of the propensity score as a linear term in the imputation model (Firth and Bennett, 1998; Bang and Robins, 2005). Another approach is to calibrate the predictions from a parametric model by adding means of the weighted residuals, with weights equal to inverse of the propensity scores (Robins, Rotnitzky and Zhao, 1994; Scharfstein, Rotnitzky and Robins, 1999). In this talk, I compare the PSPP method with these methods by simulation.

Variable Selection with Missing Data in Survival Analysis

❖ Kun Liang, Iowa State University, 428 S Walnut Ave 24, Ames, IA 50010, *liangkun@iastate.edu*; Michael D. Larsen, Iowa State University; Terry M. Therneau, Mayo Clinic

Key Words: missing data, multiple imputation, survival analysis, variable selection

Missing data bedevils medical data. Multiple imputation (MI) has been proposed and carried out on some of the previous studies, but how to do variable selection with multiple imputed dataset has not been well studied. We explored the issues implementing MI with medical data and compared several ways doing variable selection after MI in a case study of a Multiple Myeloma survival data set. With the additional help of bootstrap, we claim our final model is more reliable than complete case only method that was used before.

Which Missing Value Imputation Method To Use in Expression Profiles: A Comparative Study and Two Selection Schemes

❖ Guy Brock, University of Louisville, Bioinformatics and Biostatistics, 555 S Floyd Street, Louisville, KY 40202, *guy.brock@louisville.edu*; John Shaffer, University of Pittsburgh; Richard Blakesley, University of Pittsburgh; Meredith Lotz, University of Pittsburgh; George C. Tseng, University of Pittsburgh

Key Words: missing values, microarrays, entropy, self-training, partial least squares, k-nearest neighbors

Gene expression data frequently contain missing values, however, most downstream analyses for microarray experiments require complete data. Many methods have been proposed to estimate missing values, but the specific conditions for which each method is preferred remains largely unclear. In this report we describe an extensive evaluation of current imputation methods on multiple types of microarray experiments. We found that the success of each method depends on the underlying “complexity” of the expression data, i.e. the difficultly in mapping the gene expression matrix to a lower-dimensional subspace. We develop an entropy measure to quantify this complexity, and introduce two complementary selection schemes, entropy-based selection (EBS) and self-training selection (STS), for determining the most appropriate imputation method for any given dataset.

Interval Estimation for the Difference in Median Survival Times Based on Multiple Imputation

❖ Maria Kocherginsky, The University of Chicago, 5841 South Maryland Ave, MC 2007, Chicago, IL 60637, *mkocherg@uchicago.edu*; Theodore G. Karrison, The University of Chicago

Key Words: Survival Analysis, Difference in Medians, Confidence Interval, Multiple Imputation

In survival analysis interest is often focused on the median time to event. Generation of one-sample confidence intervals for the median failure time has received considerable attention in the literature, but less attention has been given to interval estimation of the difference in medians between two groups. We propose a nonparametric method based on multiple imputation approach. Complete survival times for censored observations are first imputed by drawing from the conditional distribution given survival up to the time of censoring. The difference in medians and its variance for the complete data are then estimated using Price and Bonett (2002), and Rubin’s multiple imputation procedure is used to obtain the total variance from which confidence intervals are derived. Simulations show reasonably accurate coverage probabilities for sample sizes of 25-75 per group and 10-40% censoring.

Analysis of Longitudinal Data with Informative Dropout

❖ Li Chen, The University of North Carolina at Chapel Hill, Department of Biostatistics, CB#7420 McGavran-Greenberg, Chapel Hill, NC 27599, *lchen@bios.unc.edu*; Danyu Lin, The University of North Carolina at Chapel Hill; Donglin Zeng, The University of North Carolina at Chapel Hill

Key Words: Counting process, Repeated measures, Informative dropout, Linear regression, Incomplete data

The analysis of longitudinal data is often complicated by informative dropout. We specify a semiparametric bivariate linear regression model for the repeated measures and the time to informative dropout. We develop efficient estimating functions for the regression parameters. The resulting

estimators are shown to be consistent and asymptotically normal with limiting covariance matrices that can be estimated through an efficient resample method. Simulation studies show that the proposed estimators perform well in practical situations and are more efficient than those of Lin and Ying (2003, Biostatistics). An application to an HIV study is provided.

Semiparametric Marginalized Model for Longitudinal Data with a Terminating Event

❖Mengling Liu, New York University, 650 First Ave, Room 526, New York, NY 10016, *mengling.liu@med.nyu.edu*; Wenbin Lu, North Carolina State University

Key Words: Generalised estimating equations, Generalised linear model, Marginalised random-effects model, Nonignorable dropout, Transformation model

We investigate marginal inference of longitudinal data when there exists a dependent terminating event. The proposed marginalized model directly specifies marginal associations between longitudinal responses and covariates and incorporates the dependent terminating event, which marginally follows a semiparametric transformation model, through a flexible conditional mean model. We develop an estimation procedure based on a series of asymptotically unbiased estimating equations. The resulting regression estimators are shown to be consistent and asymptotically normal, with a sandwich-type variance-covariance matrix that can be consistently estimated by the usual plug-in rule. The proposed approach is evaluated by simulations under practical settings and illustrated by real data applications.

469 Survival Data Analysis I

Biometrics Section

Wednesday, August 1, 2:00 pm–3:50 pm

Correlating Large-Scale Biological Data with Censored Survival Data

❖Karthik Devarajan, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111, *karthik.devarajan@fccc.edu*

Key Words: supervised learning, high-throughput study, partial least squares, gene expression, microarray, censored survival data

The advent of high-throughput technologies such as microarrays has resulted in large amounts of biological data in the form of expression profiles of thousands of genes and proteins. In recent years, there has been a tremendous interest in linking gene and protein expression data with outcome variables using supervised learning methods. An important application lies in correlating such large-scale data with censored survival data where the gene expression profile of a patient is used to predict the survival probability. In this paper, we survey the literature in this area as well as propose methods that combine learning theoretic approaches with survival models for censored data. We illustrate our methods via real-life cancer microarray data as well as simulations.

Empirical Likelihood-Based Inference for the Calibration Regression Model with Medical Cost

❖Yichuan Zhao, Georgia State University, Department of Mathematics and Statistics, Atlanta, GA 30303, *yzhao@mathstat.gsu.edu*

Medical cost has received increasing interest recently in Biostatistics. Statistical analysis of life time medical cost has been challenging by the fact that the survival times are censored on some study subjects and their subsequent cost are unknown. Huang(2002) proposed the calibration regression model which is a semiparametric regression tool to study the medical cost associated with covariates. In this talk, an inference procedure is investigated using empirical likelihood (EL) method. An adjusted EL confidence region is constructed for the vector of regression parameters. We compare the proposed EL method with normal approximation method. Simulation results show that the proposed EL method outperforms the normal approximation method in terms of coverage probability. In particular, the adjusted empirical likelihood overcomes the under coverage problem.

Martingale Residuals for Nested Case-Control Data

❖Ornulf Borgan, University of Oslo, Department of Mathematics, PO Box 1053 Blindern, Oslo, N-0316 Norway, *borgan@math.uio.no*; Bryan Langholz, University of Southern California

Key Words: Cohort sampling, Counter-matching, Counting processes, Cox's regression model, Regression diagnostic plots, Relative risk regression

Standard use of Cox's regression model and other relative risk regression models for censored survival data requires collection of covariate information on all individuals under study even when only a small fraction of them die or get diseased. For such situations nested case-control sampling offers a useful alternative. For cohort data, martingale residuals provide a useful tool for assessing the fit of a model. We introduce a natural extension of the cohort methods to nested case-control data, and show that plots of martingale residual processes provide a useful tool for checking model-fit. A formal goodness-of-fit test to go along with the plots is also presented. The methods are illustrated using data on lung cancer deaths in a cohort of uranium miners.

Power of Nonparametric Logrank and Wilcoxon Tests with Adjustment for Covariates: A Simulation Study

❖Honghua Jiang, Eli Lilly and Company, 14421 Chariots Whisper Dr, Westfield, IN 46074, *jianghh@lilly.com*; James Symanowski, Nevada Cancer Institute; Sofia Paul, Novartis Pharmaceuticals; Yongming Qu, Eli Lilly and Company; Anothy Zagar, Eli Lilly and Company; Shengyan Hong, Eli Lilly and Company

Key Words: Time-to-event variables, Cox proportional hazards model, analysis of covariance

Time-to-event variables are common primary and secondary outcomes for oncology trials. Usual methods of analysis for these endpoints involve product-limit estimators of cumulative survival for treatments, logrank tests to compare treatment groups and Cox proportional hazards model with treatment and potential prognostic covariates. Adjustment for covariates reduces bias and may increase precision and power. However, the appropriateness of Cox proportional hazards model depends on parametric assumptions. One way to address this issue is to use nonparametric analysis of covariance (Tangen & Koch 1999). Simulation studies are carried out to investigate the power of nonparametric tests with adjustment of covariates vs. nonparametric tests without adjustment of covariates. A comparison between adjusted and unadjusted methods is also illustrated with an oncology clinical trial example.

The Cost of Checking Proportional Hazards

❖Bryan Shepherd, Vanderbilt University, 1161 21st Avenue South, S2323 MCN, Nashville, TN 37232, *bryan.shepherd@vanderbilt.edu*

Key Words: survival analysis, Cox regression, bootstrap, prediction, model uncertainty, AIDS

Confidence intervals (CI) and the reported predictive ability of statistical models may be misleading if one ignores uncertainty in the model selection procedure. When analyzing time-to-event data using Cox regression, one typically checks the proportional hazards (PH) assumption and subsequently alters the model to address any violations. Such an examination and correction constitutes a model selection procedure, and if not accounted for could result in misleading CI. With the bootstrap, I study the impact of checking the PH assumption using (1) data to predict AIDS-free survival among HIV-infected patients initiating antiretroviral therapy, and (2) simulated data. CI that adjust for the PH check tend to have similar or better coverage. The impact of checking the PH assumption is greatest when the p-value of the test for PH is close to the test's chosen Type I error probability.

Nonparametric Estimation of State Occupation, Entry, and Exit Times with Current Status Data

❖ Ling Lan, University of Louisville, 555 S Floyd St Ste 4026, Bioinformatics and Biostatistics, Louisville, KY 40202, *l0lan001@louisville.edu*; Somnath Datta, University of Louisville

Key Words: Multistate model, state occupational probability, current status data, PAV, nonparametric regression, competing risks

As a type of multivariate survival data, multistate models have a wide range of applications, notably in cancer and infectious disease progression studies. We revisit the problem of estimation of state occupation, entry and exit times in a multistate model, where various estimators have been proposed in the past under a variety of parametric, semi-parametric and nonparametric assumptions. We primarily focus on two nonparametric approaches, one using a product limit formula and the other using a fractional risk set calculation, followed by a subtraction formula to calculate the state occupation probability of a transient state. A numerical comparison between the two methods is presented using extended simulation studies. We illustrate the two approaches using a breast cancer dataset extracted from the SEER cancer registry.

Assessing Fit in Cure-Mixture Models

❖ E. Paul Wileyto, University of Pennsylvania, Tobacco Use Research Center, Psychiatry, 3535 W Market Street, Suite 4100, Philadelphia, PA 19104-3309, *epw@mail.med.upenn.edu*; Daniel F. Heitjan, University of Pennsylvania

Key Words: residuals, cure models, acellerated failure time, proportional hazards, survival, behavioral medicine

Cure-models fit event times for a fraction of subjects truly at risk with a parametric survival model, and membership in the cured fraction using logistic regression. We have modified the Schoenfeld residual (from Cox regression) to assess adequacy of the parametric cure model for effects of covariates. The original residual is the difference between observed and expected covariate values for subjects failing at each event time; expectation is calculated by weighting covariate values in the risk set by proportional hazards (PH). Our alternative weights use the full hazard in place of PH, and account for the probability of membership in the cured class. We tested performance by simulation. The residuals assess fit well in both PH and accelerated failure-time models, and detect misspecification. We illustrate using relapse times from a smoking cessation study. Support: NCI/NIDA P5084718.

470 Statistical Issues in Analyzing Clinical Trials Data ● ✪

Biopharmaceutical Section

Wednesday, August 1, 2:00 pm–3:50 pm

Analysis of Cumulative Distribution of Responders in Clinical Trials

❖ Manuela Buzoianu, Johnson & Johnson PRD, 1125 Trenton-Harbourton Road, Titusville, NJ 08560, *mbuzoian@prdus.jnj.com*; Akiko Okamoto, Johnson & Johnson PRD; Surya Mohanty, Johnson & Johnson PRD; Yi-Wen Ma, Johnson & Johnson PRD

Key Words: clinical trial, responder analysis, dropout, hypothesis test, simulations

In clinical studies it is common to perform a responder analysis to compare two treatment groups. In general, this analysis consists in testing treatment effects difference using proportions of subjects showing a response, where the criterion for response is mainly based on an arbitrary cut-off. Instead of a single cut-off, the response definition can be just as well defined as a continuum of cut-offs and thereby resulting in a cumulative distribution of response, also called responder curve. Methods like log-rank, Peto-Peto, Harrington-Fleming and Kolmogorov-Smirnov nonparametric tests can be utilized towards a comparative analysis of the responder curves corresponding to different treatment groups. The performance of all the aforementioned tests will be assessed by their power, estimated using simulations under various dropout models.

Binary Clinical Trials Data Analysis Using Pairwise Likelihood

❖ Juanmei Liu, University of California, Los Angeles, 3265 S Sepulveda Blvd, Apt 309, Los Angeles, CA 90034, *jliuuanm@stat.ucla.edu*; Peter M. Bentler, University of California, Los Angeles

Key Words: SEM, polychoric correlation, polyserial correlation

This paper develops the statistical theory for a new multistage ordinal methodology for structural equation modeling, based on a recently developed maximum pairwise likelihood method. Unlike earlier methods, the maximum pairwise likelihood approach maximizes an objective function based on the product of bivariate probabilities from any two different pairs of variables to estimate both thresholds and polychoric/polyserial correlations. The asymptotic distribution of these estimators is used to develop a second stage estimation and testing procedure for SEM based on generalized least squares, and a new goodness-of-fit statistic is obtained that is asymptotically chi-square distributed. Results from a simulation study and a real example are provided to illustrate the methodology.

A Mathematical Comparison of a Randomized Withdrawal Clinical Trial Design and a Parallel Groups Design To Demonstrate Disease Modification

❖ Suzanne Hendrix, Myriad Pharmaceuticals, 320 Wakara Way, Salt Lake City, UT 84108, *shendrix@myriad.com*; Scott Horton, Myriad Pharmaceuticals; Alexander Gutin, Myriad Pharmaceuticals

Key Words: disease modification, clinical trial, randomized withdrawal, staggered start, natural history randomized start

Two crossover designs have been proposed in order to demonstrate disease modification in clinical trials: the Randomized Withdrawal and the Staggered Start studies. An effect that continues after the drug is removed is disease modifying, but this Randomized Withdrawal design is difficult to implement due to ethical concerns. A Staggered Start design avoids this issue, but requires a long trial period. These study designs are compared mathematically and are shown to be equivalent with minimal assumptions. A new 'natural history staggered start' design using parallel groups is shown under certain conditions to be equivalent to or more stringent than the crossover designs. In these conditions it is superior, because it allows measurement of the same underlying patterns of drug effect, without the operational complications of longer trials, drop-out bias, and ethical concerns.

Futility Assessment with Two Correlated Endpoints

❖Meehyung Cho, sanofi-aventis, mail box BX2 416B, 200 Crossing Blvd, Bridgewater, NJ 08807-0890, *meehyung.cho@sanofi-aventis.com*; Bill Stager, sanofi-aventis

Key Words: futility, conditional power, predictive power, correlation

Interim assessment of accumulating data from ongoing clinical trials is performed for various reasons, including stopping for futility. Several methods have been developed to assess futility, including conditional power and predictive power calculations, but these are focused on one univariate outcome variable only. However, some clinical trials may need to consider two correlated endpoints for futility assessment at the interim analysis. Such situations are: 1) trials with co-primary endpoints that are correlated, 2) futility assessment at the interim analysis based on a surrogate marker (or biomarker) which is a secondary endpoint. In this talk, we consider extensions of the existing methods of conditional power and predictive power for futility assessment to situations where two correlated endpoints should be considered for futility, and apply to a clinical trial example.

Reporting Quality-of-Life Results Using Percent Change from Baseline Analysis

❖Jen-Fue Maa, Bristol-Myers Squibb Company, 1 Colonial Ct, Plainsboro, NJ 08536, *jenfue.maa@bms.com*; Kao-Tai Tsai, Bristol-Myers Squibb Company; Daniel Seekins, Bristol-Myers Squibb Company; Rozina Khanna, Bristol-Myers Squibb Company

Key Words: FAHI, HIV/AIDS, patient reported outcome

It is often a challenge to interpret if a change in quality of life (QoL) scores is clinically significant since different measures may produce different results. A recent HIV trial demonstrated significant improvements in CD4 cell count and QoL using a psychometric instrument (FAHI) in virally suppressed patients switching HIV medications. Both percent change and absolute change from baseline in physical function of QoL were correlated with improvement in CD4 cell count (r=0.21 p=0.001 and r=0.18 p=0.004 respectively) with higher correlation observed in the percent change analysis. Absolute change does not show a definitive mathematical advantage over percent change and vice versa for correlation to clinical outcomes. Changes in absolute numbers may not be easy to interpret for most clinicians. We propose to interpret the finding by percent improvement in QoL.

Distributional Approach for Measure of Agreement

❖Weishi Yuan, Food and Drug Administration, 1401 Rockville Pike, 200 South, Rockville, MD 20852, *vivian.yuan@fda.hhs.gov*; Tie-Hua Ng, Food and Drug Administration

Key Words: diagnostic test, Measure of agreement,

A distributional approach is proposed as a measure of agreement for a continuous variable between two methods/raters so that the null hypothesis that the two methods/raters are the same may be tested. Simulated data is used to illustrate the procedure. A real example will be presented.

Test Homogeneity of Risk Difference Across Subgroups in Clinical Trials

❖Lanju Zhang, MedImmune, Inc., One Medimmune Way, Gaithersburg, MD 20878, *zhangla@medimmune.com*; Harry Yang, MedImmune, Inc.

Key Words: score test interval

A new method on test of homogeneity of risk difference across different groups in clinical trials is proposed. This method is a combination of interval estimation of risk difference based on score test and traditional weighted least squares statistics. The type I error rate of the new method should be robust against the relative small sample size. The performance of the new method will be simulated and compared to the traditional method.

471 Application of Innovative Analysis in Clinical Trials ● ✪

Biopharmaceutical Section

Wednesday, August 1, 2:00 pm–3:50 pm

Using Latent Sleepiness To Evaluate an Important Side Effect of Promethazine

❖Alan Feiveson, National Aeronautics and Space Administration, Johnson Space Center, Mail Code SK3, Houston, TX 77058, *alan.h.feiveson@nasa.gov*; Matthew Hayat, National Aeronautics and Space Administration; Zalman Vaksman, Wyeth; Lakshmi Putcha, National Aeronautics and Space Administration; Jason L. Boyd, National Aeronautics and Space Administration

Key Words: latent variable, instrumental variable, sleepiness score, drug concentration, cognitive performance

Astronauts often use promethazine (PMZ) to counteract space motion sickness; however PMZ may cause drowsiness, which might impair cognitive function. In a NASA ground study, subjects received PMZ and performed several cognitive tests during the drug metabolization period. Subjects also reported sleepiness using the Karolinska Sleepiness Score (KSS), which ranges from 1 - 9. A problem arises when using KSS to establish an association between true sleepiness and performance because KSS scores tend to overly concentrate on the values 3 ("fairly awake") and 7 ("moderately tired"). Therefore, we defined a "latent" sleepiness measure as a continuous random variable describing a subject's actual, but unobserved true state of sleepiness through time. The latent sleepiness and observed KSS were associated in a conditional probability model to permit unbiased comparison of performance measures.

The Entropy Correlation Coefficient and Its Application to Clinical Data Analysis

❖Ronald Pearson, ProSanos Corporation, 225 Market St, Suite 502, Harrisburg, PA 17101, *ronald.pearson@prosanos.com*; Jaakko Astola, Tampere University of Technology

Key Words: association measures, entropy correlation, clinical data analysis

The entropy correlation coefficient (ECC) is a measure of association between nominal variables proposed by Astola and Virtanen in 1981. Recently, Craddock et al. suggested that the ECC is useful in the exploratory analysis of clinical data and that its properties deserve further study. This paper compares the performance of the ECC with four better-known alternatives, first for a set of simulation examples and then for three clinical datasets, including the one considered by Craddock et al. The simulation examples permit quantitative control of the degree of dependence between the two sequences, independent of potentially important factors like the number of levels for each nominal variable and their marginal distributions. The three clinical examples illustrate the range of practical problems where these measures are potentially useful.

Propensity Score Application in Historically Controlled Clinical Trials

❖Daniel Reyner, Zimmer, Inc, PO Box 708, Warsaw, IN 46581-0708, *daniel.reyner@zimmer.com*

Key Words: Propensity scoring, clinical trial, randomization, historical control, prospective, retrospective

Clinical data used to achieve regulatory approval confer a population of historical controls to the study sponsor. Such rich data sources are valuable when planning a new clinical trial. Provided an adequate amount of high quality legacy data is available, such data can provide a historical control for a prospective clinical trial. The method of propensity scoring has appeal in such situations as, in principal, it leads to comparisons which approximate results which would be obtained in randomized active controlled clinical trials. This paper examines the use of propensity scoring in historically controlled clinical trials, and provides results from an application of recommended propensity scoring methods to historically controlled clinical trial data.

Improving the Delta Method

❖Paul Duty, Montgomery College, 51 Mannakee Street, Rockville, MD 20850, *paulduty@comcast.net*; Nancy Flournoy, University of Missouri-Columbia

Key Words: confidence interval estimation, coverage improvement, statistical theory, rational functions, logistic model, quantal estimation

Confidence intervals for nonlinear functions of parameters, when based on the Delta Method, seem prone to undercoverage. We propose a strategy for constructing confidence intervals that we call the Composition Method. The Composition Method is a completely analytical approach that produces asymmetric intervals that have higher coverage than the Delta Method interval for many nonlinear functions, including rational functions. We show why Composition Method intervals have higher coverage. Analytical comparisons are then made between Composition Method intervals and the Delta Method interval. Finally, we compare the coverage between confidence intervals based on the square root and logarithmic compositions with the coverage of the Delta Method interval for the quantal from a dose-response function.

The Range Disparity Distribution and Its Applications

❖Lawrence Marinucci, Acorda Therapeutics Inc, 15 Skyline Drive, Hawthorne, NY 10532, *lmarinucci@acorda.com*; Tailiang Xie, Brightech International

Key Words: Range disparity, Response analysis, Clinical trial

Conventionally, inferential statistics focus on measures of central tendency, e.g. the population mean and variance. Little attention is given to the range of sample distributions, while the range is often associated with exceptional values. Under certain circumstances, effects visible at the ranges of a sample may provide more sensitive and even more meaningful measures of differences between populations than the differences in means. In this paper we introduce a distribution of "range disparity" to describe the probability that a given number of items in one set fall outside the range of all the items in another set on a given measure. It can be used to form statistical inferences from the range differences of two independent samples, irrespective of the underlying population distribution. We will also provide several practical application examples.

Preference-Based Analysis for Estimating Treatment Efficacy Among Compliant Patients in a Randomized Clinical Trial

❖Afisi S. Ismaila, McMaster University, Clinical Epidemiology and Biostatistics, 1280 Main St. West, T-27 Building, Room 324, Hamilton, ON L8S 4L8 Canada, *ismailas@mcmaster.ca*; Stephen Walter, McMaster University

Key Words: Randomized controlled trials, compliance, efficacy, treatment preference

When there are departures from assigned treatments in a randomized controlled trial (RCT), researchers are faced with the problem of estimating treatment benefits in compliant patients while respecting randomization. Traditional approaches include intention-to-treat (ITT), per protocol (PP) and as-treated (AT) analyses. We present an extension of the preference-based (PB) analysis proposed by Walter et al. (JCE 2006) for estimating the magnitude of a treatment effect among compliers by classifying patients into 5 preference groups. Specifically, we will discuss the statistical properties of the estimators for binary and continuous outcomes. We will also compare the PB analysis with traditional analyses in the estimation of treatment efficacy in compliant patients.

Bootstrap Test for Assessing the Efficacy of Drug Combinations

❖Liwen Xi, Johnson & Johnson PRD, Welsh and McKean Roads, PO Box 776, Spring House, PA 19477-0776, *lxi2@prdus.jnj.com*; Guoyong Jiang, Cephalon, Inc.; Jianjun Li, Merck & Co., Inc.

Key Words: bootstrap-min test, min-test, composite hypothesis

In clinical trials, a new therapy is often comprised of a combination of two or more drugs. It is necessary to demonstrate that the combination therapy is more efficacious than each of its drug constituents alone. A classical statistical method for assessing the superiority of a combination therapy is the min test, whose critical value is determined conservatively. We propose the bootstrap-min test for assessing the efficacy of a combination therapy which entails using the bootstrap technique, resulting in an increase in power. In addition, this new test does not have any distribution requirements for the outcome variable of interest; therefore, it offers more flexibility compared with the classical min test. Power and type-I error rate comparisons are carried out using simulations between the new bootstrap and currently available tests.

472 Copulas, Covariances, and Risk

Business and Economics Statistics Section

Wednesday, August 1, 2:00 pm–3:50 pm

Efficient Estimation of Elliptical Copula Models

❖Mei-Mei Zen, National Cheng-Kung University, No 1 University Road, Tainan, 70101 Taiwan, *mmzen@stat.ncku.edu.tw*

Key Words: elliptical distribution, copula, MLE, IFM

Value-at-Risk (VaR) of a portfolio is determined by the multivariate distribution of risk factor increments. The RiskMetrics approach, a widely used methodology for VaR estimation, is based on the assumption of multivariate normality. In this study we suggests that the returns are elliptically distributed which better describe the empirical characteristics of financial returns. The distribution of returns will be modeled through copula approach with parameters. The statistical inference will be developed, such as the MLE (maximum likelihood estimator), IFM (Joe, 1997) etc. The asymptotic efficiency of IFM w.r.t MLE will be derived analytically. The result shows that IFM is an efficient estimator for considered models.

Bayesian Wavelet Estimation of Copulas for Dependence Modeling

❖Leming Qu, Boise State University, Department of Mathematics, 1910 University DR, Boise, ID 83725-1555, *lqu@boisestate.edu*

Key Words: Wavelet, Bayesian, Copulas, MCMC, Dependence Modeling

Copulas are full measures of dependence among random variables. It is now widely used in financial engineering for modeling high-dimensional problems, such as value-at-risk or portfolio credit risk. A copular's hidden dependence structure that couples a joint distribution with its marginals makes a parametric copular model non-trivial. We propose a nonparametric estimator using wavelet in the Bayesian framework. A mixture prior of a point mass at zero and a normal distribution is imposed on the wavelet coefficients. The Markov chain Monte Carlo (MCMC) algorithm is used for posterior inference. Performances are evaluated on simulated data and on a real dataset.

Statistical Analysis for Financial Ultra-High Frequency Data: A Filtering Approach

❖Yong Zeng, University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, *zengy@umkc.edu*

Key Words: filtering, marked point process, ultra high frequency data, Bayesian Inference

Financial ultra-high-frequency data (UHF) are naturally modeled as a marked point process. We propose a general filtering model for UHF data allowing inputs from other observable factors. The model encompasses important existing models such as Exponential ACD models. The statistical foundations of the proposed model - likelihoods, posterior, likelihood ratios and Bayes factors - are studied. They are characterized by stochastic differential equations such as filtering equations. Convergence theorems for consistent, efficient algorithms are established. Two general approaches for constructing algorithms are discussed. One approach is Kushner's Markov chain approximation method, and the other is Sequential Monte Carlo} method or particle filtering} method. Bayesian inference via filtering is studied in detail. Simulation and real stock price data examples are provided, if time permits.

The Economic Value of Intelligently Subsampled Realized Covariance Estimation of Asynchronous and Noisy High-Frequency Data

❖Lada Kyj, Rice University, PO Box 1892 MS 138, Houston, TX 77251, *ladakyj@rice.edu*; Katherine Ensor, Rice University; Barbara Ostdiek, Rice University

Key Words: Realized Covariance, Market Microstructure, Asynchronous Observations, Subsampling, Volatility timing, High-frequency Data

Assessing the economic value of portfolios constructed with more precise estimates of covariance is of great interest in finance. We present a realized covariance estimator that incorporates cross-market tick-matching and intelligent subsampling. These features of the estimator offer the potential for improved performance in the presence of asynchroneity and market microstructure noise. Specifically, tick-matching preserves information when arrival structures are nonsynchronous, intelligent sampling reduces microstructure-induced noise, and averaging reduces variance. We compare the performance of this estimator with prevailing methodologies in a simulation study and by assessing out-of-sample volatility-timing portfolio optimization strategies. Results show that our estimator has smaller MSE, smaller bias, and greater economic utility than prevailing methodologies.

Optimal Investment in a Defaultable Bond

❖Weijian Liang, New York University, 44 West 4th St, Floor 8 Room 160, New York, NY 10012, *wliang@stern.nyu.edu*; Peter Lakner, New York University

Key Words: Corporate Bond, Default Risk, Utility Maximization, Optimal Investment

The present paper analyzes the optimal investment strategy in a defaultable (corporate) bond and a money market account in a continuous time model. The treatment of information on the firm's asset value is based on an approach unifying the structural model and the reduced-form model. Specifically, the asset value will be assumed to be observable only at finitely many time points before the maturity of the bond. The optimal investment process will be worked out first for a small time-horizon with a general risk-averse utility function, then a multi-period optimal strategy with logarithmic and power utility will be presented using backward induction. The optimal investment strategy is analyzed numerically for the logarithmic utility.

Portfolio-Tracking Performance When Optimized with Estimation Error

❖Andrew Siegel, University of Washington, Business School, Box 353200, Seattle, WA 98195, *asiegel@u.washington.edu*

Key Words: Finance, Portfolio, Mean Variance Optimization, Tracking Error

The actual performance of mean-variance tracking-error-optimized portfolios is inferior to the anticipated performance when the optimization is performed in the presence of estimation error, due to systematic bias that occurs when the optimization procedure exploits the noise in the data. Given a universe of n assets, a benchmark portfolio weight, and a target mean, we wish to find the portfolio with this target mean that minimizes the variance of its deviations from the benchmark. When the asset mean vector and covariance matrix are known, the optimal portfolio can be determined. When estimated means and covariances are used instead, the resulting portfolio will achieve neither the target mean nor the anticipated error variance. We use matrix Taylor series expansions with respect to the estimation error in order to derive second-order approximations to these systematic biases.

Biased Bootstrap Methods for Semiparametric Models

❖Mihal Giurcanu, University of Florida, 103 Griffin/Floyd Hall - P.O. Box 118545, Gainesville, FL 32603, *giurcanu@stat.ufl.edu*

Key Words: Biased Bootstrap, Block Bootstrap, Recycling, Time Series, Bias Reduction, Optimal Block Size Selection

A new biased bootstrap methodology for moment condition models has been developed. This biased bootstrap is a form of weighted bootstrap with the weights chosen to satisfy the constraints imposed by the model. First, a least favorable parametric family of weighted empirical distributions associated with the model is constructed. By resampling from this family, we "mimic" the parametric bootstrap for semiparametric models. Using the recycling algorithm, an iterated biased bootstrap is proposed, which is computationally feasible and has a higher degree of accuracy. An extension of this methodology for dependent observations is also presented. Some consistency results, applications to bias reduction and optimal block size selection are provided.

473 Investing Optimally ✪

Business and Economics Statistics Section
Wednesday, August 1, 2:00 pm–3:50 pm

Competitive Inventory Management in Treasury Markets

❖Rohan Christie-David, University of Louisville, 2814 Shoemaker Dr, Louisville, KY 40241, *rohan.christiedavid@louisville.edu*; William T. Moore, University of South Carolina; Arjun Chatrath, University of Portland

Key Words: Treasury, Bid, Ask, Spreads, Inventory, Management

Using on-the-run two-year and 10-year Treasury Notes we test four hypotheses about inventory costs and dealers' quote-setting behavior. First we decompose the bid-ask spread and then examine whether dealers in this market make larger adjustments in the inventory component (1) following macroeconomic news events, (2) towards the start and end of New York trading hours, (3) before open market operations by the Federal Reserve, and (4) when transaction sizes are relatively large. Our findings provide support for all four hypotheses.

Time-Varying Coefficient Model for Corporate Bond Credit Spread

❖Yuejiao Ma, The Pennsylvania State University, 720 North Allen Street, State College, PA 16803, *hma@stat.psu.edu*

Key Words: Time varying Coefficient Model, Markov Chain Monte Carlo, hierarchical model, credit spreads

We propose to examine potential determinants of credit spreads on corporate bonds. We consider two approaches in our analysis. The approach we consider is to combine the time varying coefficient model and hierarchical model for the prices of corporate bonds in a continuous time setting, then estimate the model using the Markov Chain Monte Carlo method. Data used for this analysis include daily prices and yields for all straight bonds form May 28, 2002, to June 10, 2005.

Trading American Stock Options Amidst Market/Model Disparity

❖Scott Nickleach, University of Pittsburgh, Rd2 Box26, Worthington, PA 16262, *snickleach@yahoo.com*

Key Words: Monte Carlo, stock option, simulation

A disparity between model prices and market prices is a common shortcoming of most stock option pricing models. We shall consider the problem of trading American style exchange traded options under certain model mispricing. We use a series of estimated conditional expectations to formulate trading strategies, with the aim of gaining some level of control over the distribution of returns.

Testing the Informational Efficiency of Municipal Bond Market

❖Lei Zhang, Syracuse University, 721 University Ave Suite 500, Syracuse, NY 13244, *lzhang06@syr.edu*; Chunchi Wu, Syracuse University; Raja Velu, Syracuse University

Key Words: Generalized Spectral Test, Nonlinear Dynamics, Martingale, Real-time Transaction Reporting System, Market Efficiency, Municipal Bond Index

Municipal bond market has been known for lack of transparency. However, the Real-time Transaction Reporting System, launched in 2005, replacing the daily reporting rule, requires all the transactions to be reported within 15-minutes of completion. This big move toward the timely reporting system brought the municipal bond market efficiency issue to our attention because with more "real-time" trading information available, the municipal market is expected to become more efficient. In this essay we want to see how this change influences the municipal market return dynamics. Since the available tests of market efficiency have no power against the alternative of nonlinear dependence with zero-correlation, we test both linear and nonlinear dependence of the CBOT Bond Buyer 40 Municipal Bond Index return via Hong's (1999) Generalized Spectral density test.

Principal Analysis and Price Dynamics for Sparse Auction Data

❖Bitao Liu, University of California, Davis, Department of Statistics, One Shields Ave., Davis, CA 95616, *btliu@wald.ucdavis.edu*; Hans-Georg Müller, University of California, Davis

Key Words: eBay Auctions, Derivative Estimation, Functional Data Analysis, Principal Component, PACE

Observed bids from eBay auctions are usually sparse in the middle and more frequent near the beginning and especially near the end of the auction. For auctions related to identical products, one may view the bid history of each of these auctions as realization of an underlying smooth stochastic process. It is then of interest to recover the price curves and also their derivatives. We demonstrate the application of the principal analysis through conditional expectation (PACE) approach for such data. This functional data analysis method does not require pre-smoothing and uses functional principal component analysis for sparse data. We propose the extension of this approach to the estimation of derivatives, where sparse data are particularly problematic, aiming at exploring auction price dynamics.

Empirically Confronting Stochastic Singularity: An Application to the Cox, Ingersoll, and Ross Model

❖Kenneth Roskelley, The University of Arizona, 315G MCLND, Tucson, AZ 85722, *kdr@email.arizona.edu*; Christopher Lamoureux, The University of Arizona

Key Words: Testing Models of Arbitrage, Forecasting Interest Rates

We use encompassing regressions to evaluate interest rate forecasts along the yield curve over the period 1989 through 2005. Inputs to the encompassing regression include the forecast from one- and two-factor Cox, Ingersoll, and Ross (CIR) models–obtained by estimating the model without adding pricing errors, the slope of the yield curve, and the change in Fed target rates. At the short end, we also include the forward rate and forecasts derived from popular non-parametric models. This empirical design has three advantages. First, we maintain the stochastic singularity common to all no-arbitrage models. Second, rather than accept or reject the model's cross-sectional restrictions, we evaluate if the model contains any useful information about yield dynamics. Third, we restrict our attention to a time period of transparent Federal Reserve policy.

474 Regression and Forecasting ● ✪

Business and Economics Statistics Section
Wednesday, August 1, 2:00 pm–3:50 pm

Testing the Equivalence of Regressions for the Capital Asset Pricing Models

❖Hubert J. Chen, National Cheng Kung University, 1 University Road, Tainan, 70101 Taiwan, *hjpcheney@yahoo.com*; Jevons Lee, Zhejiang University; Li-Wei You, National Cheng Kung University

Key Words: equivalence, studentized range, mutual funds, level and power, portfolios

A studentized range statistic is proposed for testing the hypothesis of equivalence of alpha intercepts and beta slopes for the capital asset pricing regression models against any given alternative hypothesis of inequivalence. Both the level and the power of the proposed test associated with these hypotheses are obtainable at their corresponding least favorable mean configurations (LFMC). It can be seen that the level and the power at their LFMCs are completely independent of the unknown parameters. Therefore, for a given level and a given power, the critical value and the required sample size for the test can be simultaneously determined. Application to a real example on US largest mutual funds is demonstrated.

Inferences in Regression with Correlated Residuals

❖ Yue Fang, University of Oregon, CEIBS, College of Business, Eugene, OR 97403, *yfang@lcbmail.uoregon.edu*

Key Words: Autocorrelation, Hypothesis test, Generalized least squares, Regression

In this article we show that autoregressive corrections based on generalized least squares (GLS) estimation yields estimates for serially correlated regression parameters that are asymptotically efficient, and that have the same limiting distribution as those of Aitken estimators (GLS). Based on exhaustive Monte Carlo studies we also demonstrate that for finite samples the relative efficiency this estimator is comparable to that of estimated GLS when the autoregressive order is appropriately chosen. Moreover, this type of correction can yield as good, and more often, better forecasts than those generated from ordinary least squares (OLS) or from GLS using the correct form of the residual autocorrelation structure. This suggests that there is not much to be gained in trying to identify the correct form and order of the serial correlation.

Performance Comparison of Simple Forecasting Methods for Macroeconomic Indicators

❖ Bogong T. Li, Bank of America, 100 N Tryon St, NC1-007-06-11, Charlotte, NC 28255, *tim.li@bankofamerica.com*

Key Words: Economic indicators, Forecast, Exponential smoothing, Stepwise regression, Holt-Winters

Monthly economic indicators released by U.S. federal government impact financial markets the most when the direction of their change is different from market consensus. For those who have anticipated the difference will achieve significant financial gain. Given current quantitative forecast methods based on statistical models minimize only squared errors, irrespective to directional differences, optimal MSE models do not necessary produce the best forecasts. In this research we will evaluate the relationship between MSE types of criteria with directional correction using forecasts for some important U.S. macroeconomic monthly releases. As a result from this evaluation, we further propose an improved forecasting method that will perform better in forecasting direction than current methods.

A Statistical Model of a Firm's Expenditure

❖ Jonathan Hosking, IBM T. J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598, *hosking@watson.ibm.com*; Bibhas Chakraborty, University of Michigan

Key Words: business, finance, forecasting, point process

We present a statistical model of the expenditure over time of a firm or other business organization. Assume that the firm "commits" to expenditure by issuing purchase orders for goods or services from outside suppliers. Each purchase order results in one or more payments being made to the supplier: these payment events cause expenditure to be recognized in the firm's accounts. We model expenditure arising from a single commit as a marked point process with parameters that are functions of the attributes of the commit. We model the dates and amounts of future commits, using a regression approach that relates rate of commit to time of year. With these models and a schedule of future releases, we construct forecasts of future expenditure. These can be used by the firm's financial officers to manage expenditure. The models are applied to the capital expenditure of an IBM business unit.

A Comparison of Inference Procedures for the Simple Linear Errors-in-Variables Regression Model

❖ Stephen M. Miller, Bureau of Labor Statistics, 2 Massachusetts Ave. N.E., Washington, DC 20212, *miller_s@bls.gov*

We consider the simple linear errors-in-variables regression model, and compare various inference procedures available for the parameters of the model. Fuller (1987) and Carroll, Rupert, Stefanski and Craniceanu (2006) provide a survey of many of the existing methods. We look at many of these as well as some additional methods, and compare how they perform both in terms of their theoretical properties and a their performance in large simulation studies. Strengths and weaknesses of the various procedures are examined.

Multiple Imputation of Right-Censored Wages in the German IAB-Employment Register Considering Heteroscedasticity

❖ Thomas Buettner, Institute for Employment Research, Regensburger Str 104, Nuremberg, 90478 Germany, *thomas.buettner@iab.de*; Susanne Raessler, Institute for Employment Research

Key Words: multiple imputation, missing data, censored wage data, simulation study

In many large data sets of economic interest, some variables, as wages, are right-censored. In order to analyze wages with the IAB-employment register we first have to solve the problem of censored wages at the upper limit of the social security system. We treat this problem as a missing data problem and use multiple imputation approaches to impute the censored wages by draws of a random variable from a truncated distribution, based on Markov chain Monte Carlo technique. In general, the dispersion of income is smaller in lower wage categories than in higher and the assumption of homoscedasticity in an imputation model is highly questionable. Therefore, we suggest a new multiple imputation method which does not presume homoscedasticity of the residuals. Finally, in a simulation study different imputation approaches are compared to confirm the necessity and validity of the new approach.

Measuring Liquidity Costs in Agricultural Futures Markets: Conventional and Bayesian Approaches

❖ Julieta Frank, University of Illinois at Urbana-Champaign, 1952 Orchard St Apt C, Urbana, IL 61801, *julietafrank@yahoo.com*; Philip Garcia, University of Illinois at Urbana-Champaign

Key Words: agricultural futures markets, liquidity, bid-ask spread, Bayesian estimation

Estimation of liquidity costs in agricultural futures markets is challenging because bid-ask spreads are usually not observed. Spread estimators that use transaction data are available, but little agreement exists on their relative accuracy and performance. We evaluate four conventional and a recently proposed Bayesian estimators using simulated data based on Roll's standard liquidity cost model. The Bayesian estimator tracks Roll's model relatively well except when the level of noise in the market is large. We derive an improved estimator that seems to have a higher performance even under high levels of noise which is common in agricultural futures markets. We also compute liquidity costs using data for hogs and cattle futures contracts trading on the Chicago Mercantile Exchange. The results obtained for market data are in line with the findings using simulated data.

475 MCMC and Gibbs Sampling ✪

IMS

Wednesday, August 1, 2:00 pm–3:50 pm

Using Subsampling in Markov Chain Monte Carlo

❖ James Flegal, The University of Minnesota, 1901 Emerson Ave S 104,
Minneapolis, MN 55403, *jflegal@stat.umn.edu*

Key Words: Markov chain, subsampling, fixed-width, batch means

Markov chain Monte Carlo is a method of producing a correlated sample to estimate characteristics of a target distribution. Several research papers deal with the special case in estimating ergodic averages and their corresponding standard errors. A fundamental question is how to estimate the standard error of the qth quantile of π? To this end, we will use subsampling as a method of estimating the MCSE of the estimate. Further, we will compare our results to some common alternatives.

Gibbs Sampling for Bayesian Models

❖ Alicia Johnson, The University of Minnesota, 2014 DuPont Ave S Apt 7,
Minneapolis, MN 55405, *ajohnson@stat.umn.edu*

Key Words: Block Gibbs sampler, burn-in, geometric ergodicity, drift condition, minorization condition, convergence rate

We consider a Bayesian hierarchical version of the random intercepts model. The posterior density for this model makes it difficult to find closed form solutions of integrals required for inference. For this reason, we construct a block Gibbs sampler for this model. We establish drift and minorization conditions on the underlying Markov chain and apply results from J. S. Rosenthal [J. Amer. Statist. Assoc. 90 (1995) 558-566] to construct upper bounds for the total variation distance to stationarity for the chain. Using these bounds and a prespecified desired distance to stationarity we can ensure adequate burn-in for the block Gibbs sampler.

Markovian Dynamics in a RAM with Applications to Classification

❖ Zach Dietz, Tulane University, 5420 Garfield Street, New Orleans, LA
70115, *zdietz@tulane.edu*

Key Words: RAM, GEM distribution, Dirichlet distribution

The residual allocation model (RAM) constructed from a collection of iid Beta(1,theta) random variables is known as the GEM(theta) distribution. A well-known property of the GEM(theta) distribution is that the sum of the elementwise product of a GEM(theta)'s coordinates with an iid collection of discrete random variables converges to a Dirichlet distribution. Based on the work of Dietz and Sethuraman, we discuss a similar result in which the collection of discrete random variables are Markovian, and in this context we introduce a distribution on the infinite simplex that follows the recipe for a RAM, but utilizes dependent Beta random variables, and thereby giving rise to, perhaps, a nonstationary GEM distribution.

Goodness-of-Fit Tests for Generalized Linear Models When Some Covariates Are Missing

❖ Lei Jin, Texas A&M University, Department of Statistics, 3143, College
Station, TX 77843, *ljin@stat.tamu.edu*; Suojin Wang, Texas A&M
University

Key Words: Goodness-of-fit, Missing at Random, Generalized Score Statistics, Nuisance Parameters, Weighted Estimating Equations

The goodness-of-fit tests for complete samples may encounter difficulties in the presence of missing covariates. We propose goodness-of-fit tests for testing generalized linear models when some covariates are partially missing based on the generalized score statistics. The methods are based on the weighted estimating equations. The effect of nuisance parameters on the generalized score statistic are studied, including the efficiency. The asymptotic distribution of the test statistic is given. Simulation studies are also given in both ordinary linear models and logistic regressions to demonstrate the finite sample performances of this method.

476 New Developments in Spatial Modeling ●

Section on Bayesian Statistical Science

Wednesday, August 1, 2:00 pm–3:50 pm

Point Estimates for Variance-Structure Parameters in Bayesian Analysis of Hierarchical Models

❖ Yi He, sanofi-aventis, 2304 Deer Creek Dr, Plainsboro, NJ 08536,
hydinghua@gmail.com; James S. Hodges, The University of Minnesota

Key Words: Hierarchical model, point estimate, posterior distribution, precision, standard deviation, variance

MCMC has made Bayesian analysis feasible for complex hierarchical models, but the literature about variance parameters is somewhat sparse. This paper uses simulation experiments to compare three point estimators, the posterior mode, median, and mean, for variance-structure parameters in hierarchical models. The comparisons are made for three parameterizations of the variance structure, in terms of precisions, standard deviations, and variances. We first compare linear regression, where fairly explicit expressions are possible, then three more complex models: the crossed random effects model, the conditional autoregressive model (CAR) with two neighbor relations, and smoothed ANOVA. We illustrate the results using periodontal data. The surprising finding is that the posterior mode is, on the whole, the best point estimator regardless of the variance structure's parameterization.

Smoothed ANOVA with Spatial Effects as a Competitor to MCAR

❖ Yufen Zhang, The University of Minnesota, 1042 F 27th Ave SE,
Minneapolis, MN 55414, *yufenz@biostat.umn.edu*; Hodges S. James, The
University of Minnesota

Key Words: Analysis of variance, Conditional autoregressive model, Hierarchical model, Degree of freedom

Smoothed ANOVA (SANOVA; Hodges et al 2007 Technometrics) is a way to smooth ANOVA that shrinks interactions by embedding SANOVA in a hierarchical model to do shrinkage. Instead of simply shrinking effects without any structure, SANOVA can use spatial structure to smooth effects. In this talk, we will extend SANOVA to model cases in which one factor is a spatial lattice, which is smoothed using a conditional autoregressive model (CAR), and a second factor is, for example, type of cancer. As such, SANOVA may be a competitor to the multivariate CAR (MCAR) model. Simulations are done to compare SANOVA under different design matrix settings versus MCAR under different prior settings, with advantages of each approach discussed. A cancer-surveillance dataset, describing incidence of 3 cancers in Minnesota's 87 counties, is analyzed using different methods and their results are discussed.

Bayesian Spatial Models with Repeated Measurements

❖ Xiaoqian Sun, Clemson University, O104 Martin Hall, Box 340975, Clemson, SC 29634-0975, *xsun@clemson.edu*

Key Words: Spatial model, repeated measurement, slice sampler

We propose a new spatial model that takes account of the special data structure and treats a cluster of measurements as repeated measurements in one location. The model is applied to the analysis of the total vegetation coverage data in the Missouri Ozark Forest Ecosystem Project. An MCMC algorithm based on the shrinkage slice sampler is developed. The results show that the soil depth is an important factor while the aspect class is less important when modeling the total vegetation coverage. In addition, the strong spatial effect does exist in the data and four measurements in quadrants of a subplot are not strongly correlated but are not independent. Prediction of the total vegetation coverage at unmeasured locations is developed. Finally, possible generalizations are discussed.

Loglinear Modeling for Point-Referenced Spatial Data

❖ Eric Tassone, Duke University, 212 N Duke St, CEHI, Durham, NC 27701, *eric.tassone@duke.edu*; Marie Lynn Miranda, Duke University; Alan E. Gelfand, Duke University

Key Words: Bayesian disease mapping, hierarchical/multilevel models, spatial loglinear models

We develop spatial loglinear models for point-referenced data, extending nonspatial approaches by modeling a collection of spatially dependent trials of individual locations, rather than independent trials that ignore location or use only areal unit information. We use a spatial process of loglinear models, where the sampling model is a single multinomial trial at each of n (finite) locations. Despite only one observation at each location, we estimate parameters of the driving process model and predict the loglinear model at every location in the study region; the specification of the process model enables this via spatial smoothing of nearby loglinear model parameters. Further, the loglinear modeling allows point-referenced inference about all marginal and conditional probabilities associated with the model. We illustrate our approach with North Carolina Detailed Birth Record data.

Bayesian Modeling of Semicontinuous Spatial Data

❖ Virginia Recta, Food and Drug Administration, 7500 Standish Place, CVM, Rockville, MD 20855, *virginia.recta@fda.hhs.gov*; James L. Rosenberger, The Pennsylvania State University; Murali Haran, The Pennsylvania State University

Key Words: Semicontinuous variables, spatial prediction, geostatistics, generalized linear mixed models

We consider the problem of modeling realizations of a semicontinuous response Y(x) at sampling locations x. A semicontinuous response typically has a portion of responses equal to zero and a continuous distribution for the positive values. Following Diggle et al. (1998), conditional on an unobserved stationary zero-mean Gaussian process S(x), Y(x) forms a generalized linear model with S(x) appearing as an offset term in the linear predictor. We break down the response into an incidence part U and a positive part V, then specify a two-part model, $U \sim f\{u|S(x)\}$ and $V \sim g\{v|Z(x)\}$, where S(x) and Z(x) are separate but related zero-mean spatial processes with a given cross-covariance structure. We describe a new and improved fully Bayesian inferential approach for this model, focusing in particular on the choice of priors and sampling algorithms.

Sparsity and Smoothness for Disease Rate Mapping via Spatial Bayesian LASSO

❖ Haoda Fu, University of Wisconsin-Madison, 502 Eagle Heights Apt. K, Madison, 53705, *fuhaoda@stat.wisc.edu*

Key Words: Disease mapping, LASSO, Local resampling, Markov chain Monte Carlo, Markov random field, Trust region reflective Newton method

Maps of regional disease rates are useful tools for examining spatial patterns of disease. In this paper, we adopt a Bayesian approach for disease rate mapping and develop a new Markov random field prior. This prior has two hyperparameters which control the sparsity and smoothness of the estimated maps. To choose the hyperparameters, we discuss methods from both Bayesian and non-Bayesian views and propose a local resampling method for automatic selection. The maximum a posteriori (MAP) estimate is used as a summary of the posterior density function and we propose a new algorithm to calculate it efficiently. This new algorithm is more generally applicable than existing LASSO algorithms, such as LARs (see Efron et al. 2004), in terms of the ability to handle more general loss functions. This is a joint work with Professor Murray K. Clayton.

Modeling Wind Fields Using Process Convolutions

❖ Margaret Short, University of Alaska Fairbanks, P.O. Box 750125, Fairbanks, AK 99775, *jsm@mshort.authorized.yon.net*

Key Words: wind field, spatial process, process convolutions, Markov chain Monte Carlo

We propose an approach to modeling wind field data using process convolutions. Our approach has several advantages over a standard Gaussian process approach for such data. In particular, it has the ability to handle both the periodicity of the wind direction and the non-negativity of the wind speed. Ours is a Bayesian model; we implement it via Markov chain Monte Carlo. We illustrate the approach with an application.

477 Methods for Administrative and National Survey Data ● ✪

Section on Health Policy Statistics

Wednesday, August 1, 2:00 pm–3:50 pm

Evaluation of Consumer Panel Survey Data for Public Health Communication Planning: An Analysis of Annual Survey Data from 1995–2006

❖ William Pollard, Centers for Disease Control and Prevention, 2554 Circlewood Rd NE, Atlanta, GA 30345, *bdp4@cdc.gov*

Key Words: Public health, Health surveys, Health communication

Consumer survey data can provide useful information for audience analysis in public health communication planning. A widely used methodology for conducting consumer surveys involves sampling from panels of individuals and households that have been prerecruited to participate in surveys and other market research. Market research firms maintain panels of several hundred thousand or more individuals across a full range of demographics for such research. In this presentation, data from the annual national Porter Novelli HealthStyles surveys, which employ this methodology, from 1995 to 2006 with 3,000 to 4,000 respondents per year are examined. Comparisons are made with other health survey data from these same years based on national probability sampling methods. The literature on the use of panels is referenced and the implications for public health communication are discussed.

Statistical Modeling of Longitudinal Mental Distress Among Rural Seniors Who Participated in the National Population Health Survey

❖Chandima Karunanayake, University of Saskatchewan, Canadian Centre for Health and Safty in Agriculture, 103 Hospital Drive, Saskatoon, SK S7W 0W8 Canada, *cpk646@mail.usask.ca*; Punam Pahwa, University of Saskatchewan

Key Words: Mental Health, Rural Seniors, Longitudinal data, National Population Health Survey, Generalized estimating equations, Bootstrap weights

Since 1994/95 National Population Health Survey (NPHS) have been collecting longitudinal data on the physical and mental health of Canadians. The objectives were to investigate the long-term effect of addictions on mental health among seniors (55 years and older) who live in rural and urban areas and to examine the effect of risk factors such as alcohol use, smoking and excessive drug use on mental health. From the relatively wide range of mental health indicators available in the NPHS, the mental distress measure was examined as binary outcome (categories: no/low and moderate/high). The analysis was conducted using generalized estimating equation approach accounting for the complexity of multi-stage survey design using bootstrap weights available for incomplete longitudinal data. The missing data issue was also addressed.

Statistical Methods for Constructing Canadian Growth Curves

❖Stephane Tremblay, Statistics Canada, 100 Tunney Driveway, Ottawa, ON, K1A 0T6, Canada, *stephane.tremblay@statcan.ca*

Key Words: Growth curves, BMI, LMS method, Sampling weights, Design effect

Growth curves are used by Health Care providers to determine if the growth of a child or fetus (for example) is within normal boundaries. The current charts used in Canada for height, weight and body mass index (BMI) are based on United States data. Child growth curves can now be developed based on the most recent cross-sectional data available in Canada. The presentation will describe the LMS (Lambda-Mu-Sigma) method using a Box-Cox normal distribution and a cubic spline smoothing technique to estimate Canadian growth curves for children aged 2 to 18. Special attention will be put on how we used auxiliary information (from the US) to handle edge effects, how we handled extreme sampling weights and how we adjust for the design effects when estimating such curves. Criteria for the selection of the best model using statistical and graphical diagnostic tools will also presented.

Evaluation of New Programs in State Medicaid Agencies

❖Mark Carpenter, Auburn University, Department of Mathematics and Statistics, Auburn, AL 36849-5310, *carpedm@auburn.edu*

Key Words: data mining, forcasting, predictive modeling, discrimination

When state Medicaid agencies propose new programs for cost-savings or disease managements, they often have to demonstrate to the Centers for Medicare and Medicaid Services (CMS) that the programs do not increase costs (budget neutrality). This issue can be quite problematic because various programs such as preferred drug lists and others are implemented concurrently. In this presentation, we focus on one particular program in the Alabama State Medicaid Agency called Patient 1st, which currently has 420,000 Alabamians enrolled. Methods for estimating cost savings and savings sharing for participating physicians are presented. In addition to costs savings issues, we also present our attempts to characterize those eligibles who have not yet signed up for the mandatory program.

Classification Models for Identifying Chronic Disease Cases and Noncases in Administrative Claims Data

❖Lisa M. Lix, University of Manitoba, 408 727 McDermot Avenue, Manitoba Centre for Health Policy, Winnipeg, MB R3E 3P5 Canada, *lisa_lix@cpe.umanitoba.ca*; William D. Leslie, University of Manitoba; Marina Yogendran, Manitoba Centre for Health Policy; Richard Baumgartner, National Research Council Canada; Christopher Bowman, National Research Council Canada; Souradet Shaw, University of Manitoba

Key Words: prevalence, classification tree, artifical neural network, osteoporosis, administrative claims

Administrative claims data are a popular tool for chronic disease surveillance. Algorithms to identify disease cases are often based on a simple heuristic applied to individual databases. This research compares artificial neural networks and classification trees to distinguish disease cases from non-cases in multiple databases, using osteoporosis as an example. Training and test data were obtained from Manitoba, Canada. Data features included disease diagnosis codes, prescription records for disease treatment, diagnosis and prescription records for comorbid conditions, and demographics. Both classification models produced similar estimates of discriminative power and prediction error. Misclassification rates ranged from 0.31 to 0.22 depending on the data features included in the model. The model predicted probabilities were used to estimate to estimate osteoporosis prevalence.

The Minimum Level of Sensitivity and Predictive Value Positive in the Evaluation of Health Outcomes with Administrative Data

❖Yong-Fang Kuo, The University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555-0460, *yokuo@utmb.edu*; Tracy Nguyen-Oghalai, The University of Texas Medical Branch

Key Words: Sensitivity, Predict Value Positive, Measurement Error, Claims

Health administrative data is widely used to assess outcomes in population-based studies. To measure outcomes, claims based algorithms reaching sensitivity (Sen) of 90% and PVP of 70% were developed. To understand the impact of Sen and PVP on a bivariate association with dichotomous variables, a series of 100 simulations were performed for the outcome with combinations of variety levels of Sen and PVP, and different sample sizes and OR. Our results suggest that PVP of 70% can be useful in detecting OR of 1.5 with sample size of 1000. Increasing in OR or sample size increase bias and decrease coverage probability (CP), given the same PVP and Sen. PVP is more important than Sen in estimating association (by bias, mean square error, CP). The true association was always underestimated. Simulations can be used to evaluate outcome algorithms from claims, based on minimum levels of Sen and PVP.

The Impact of Deidentification on Statistical Information in Limited Access Datasets

❖April Slee, Axio Research, 2601 4th Avenue Suite 200, Seattle, WA 98121, *aprils@axioresearch.com*; Paula Diehr, University of Washington; Kenneth Rice, University of Washington

Key Words: HIPAA, Limited Access Data Set

HIPAA guidelines focus on removal of Private Health Information (PHI), or “deidentification,” when creating limited access data sets (LADS). LADS are used to estimate prevalence, calculate sample size, compare event rates and gather other information. Removing PHI impacts the amount of information available in LADS. We explored methods of quantifying the loss of information due to deidentification. We selected 14 published electronic data sets, recalculated the primary analysis, removed all PHI, and repeated the analysis on the deidentified data. We compared the ratio of confidence intervals for the predictor of interest from the original and deidentified data. The mean ratio of confidence intervals was 0.80 (p-value 0.039). Information loss was significantly greater for observational studies than randomized clinical trials (mean ratio 0.65 versus 1.0, p= 0.046).

478 Financial Data Mining and Risk ● ✪

Section on Statistical Computing, Business and Economics Statistics Section, Society for Risk Analysis, Section on Statistical Graphics, Section on Nonparametric Statistics

Wednesday, August 1, 2:00 pm–3:50 pm

Statistical Methods for Detecting Financial Crimes: Applications and Challenges

❖ Agus Sudjianto, Bank of America, 201 N Tryon, NC1-022-03-01, Charlotte, NC 28255, *agus.sudjianto@bankofamerica.com*; Sheela Nair, University of California, Los Angeles

Key Words: Financial crime, fraud, anti-money laundering, machine learning

Financial crimes affect millions of people every year, and financial institutions must employ methods to protect themselves and their customers. Fraud and money laundering are two common types of crimes, and there has been extensive research to develop algorithms to detect these crimes. Detection algorithms face a large set of challenges, including prediction of rare events, overlap between classes, and criminals who alter their strategy to trick current detection techniques. We will discuss some existing methodologies in practice as well as some challenges to stimulate further research.

An Efficient Sequential Design Method for Detecting Money Laundering

❖ Xinwei Deng, Georgia Institute of Technology, 765 Ferst Drive NW, Atlanta, GA 30332, *xdeng@isye.gatech.edu*; Roshan Joseph, Georgia Institute of Technology; Agus Sudjianto, Bank of America; Chien-Fu Jeff Wu, Georgia Institute of Technology

Key Words: Sequential design, Logistic models, Anti-money laundering

Anti-money laundering is an important responsibility for financial institutions. There are millions of transactions happening each day. Investigating each of the transactions and classifying an account to be suspicious or not is time-consuming and tedious. In this article, we propose a Bayesian sequential design method to detect money-laundering accounts. The sequential nature of the method helps to identify the suspicious accounts with minimal time and effort. The method uses logistic models and a combination of c- and D- optimality criteria to judiciously select the accounts. An application to real banking data is used to demonstrate the performance of the proposed method. Simulation study shows the efficiency and accuracy of the proposed method, as well as its robustness to model assumptions.

Unsupervised Quantile Learning To Identify Principal Direction of Anomaly

❖ Aijun Zhang, University of Michigan, 439 West Hall, 1085 South Univ Ave, Ann Arbor, MI 48109, *daerhuazhi@gmail.com*; Agus Sudjianto, Bank of America; Ming Yuan, Georgia Institute of Technology

Key Words: Unsupervised learning, Quantile, Outlier detection

We propose a quantile learning framework to tackle some unsupervised tasks. It is formulated as a maximin problem with double convex objectives. This paper studies the extreme quantile case, in order to identify the principal direction of anomaly and detect outliers. It is shown that our outlier peeling algorithm can be viewed as a sequential variate of the classical Mahalanobis depth approach.

Rule and Rule Ensemble-Based Methods for Financial Crime Detection

❖ Fernando Cela Diaz, Bank of America, 200 College St, NC1 014 13 30, Charlotte, NC 28255, *fernando.cela-diaz@bankofamerica.com*; Agus Sudjianto, Bank of America

Key Words: Financial Crime, Fraud, Machine Learning, Data Mining, Rule, Rule Ensembles

Rule inference methods are commonly used in Financial Crime Detection. Both supervised and unsupervised methods are possible. Recently, the idea of rule ensembles has been proposed as a method for improving the performance of weak rule learners. This paper will present existing methodologies and explore current real-world applications with emphasis on empirical performance evaluation in real settings.

Feature Extraction and Classification to Identify Suspicious Activities

❖ Alireza Darvish, The University of North Carolina at Charlotte, 1600 Arlyn Cir, Apt M, Charlotte, NC 28213, *adarvish@uncc.edu*; Kayvan Najarian, The University of North Carolina at Charlotte

Key Words: Feature Extraction, FFT, DWT, Money Service Business, Classification

This paper introduces a set of signal processing features that identify customers suspicious of being involved in money service business (MSB) without a license. We extract features from customer's transaction history both in time-domain or transform-domain. Time-domain features, such as the average time separation between two consecutive check transactions, can be perceived and interpreted by human experts. However, transform domain features, such as Fourier and wavelet based features, proved to extract more important features for classification. In particular, our results proved that the features extracted using wavelet transform can better identify the MSB-like behaviors. Test of statistical significance on every extracted feature using t-test and ANOVA proved the statistical significance of the extracted measures.

Simulating Goodness-of-Fit: Ensuring Fair Online Auctions

Andrew Ostarello, Exponent, Inc.; ❖ Michael Orkin, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *morkin@exponent.com*

Key Words: Simulation, R, Auctions, Goodness-Of-Fit, Cheating

In competitive financial auctions (where lenders bid on interest rates for loans, for example), there are often multiple bidders tied for the lowest rate allowed by law. In these situations, the auction winner is chosen at random from amongst the tied bidders. Given a large list of auctions, bids, and winners, the distribution of wins across the bidders takes on a complex multinomial character, making traditional goodness-of-fit measures difficult to obtain. Using parallel simulation, we attempted to determine if the tie-breaking mechanism was consistent with random selection, whether the auction was "fair", how many bidders the auctioneer can expect to be treated "unfairly" by each series of auctions, and effective methods for the auctioneer to "cheat" without being caught by our analyses.

The Netflix Challenge for Statisticians

❖J. Brian Gray, The University of Alabama, Dept of ISM, 300 Alston Hall, Tuscaloosa, AL 35487-0226, *bgray@cba.ua.edu*

Key Words: data mining, machine learning, massive data set, predictive modeling, recommendation system

In October 2006, Netflix, the online DVD rental service, announced a contest to award $1,000,000 to the first person or team to improve on the accuracy of the Netflix customer ratings prediction model by at least 10%. Netflix made available data on 100 million customer DVD ratings for the contest. Within three months, the leaders in the contest had reached 6.67% improvement. While this contest would appear to be an excellent opportunity for statisticians to prove their mettle in data mining, statisticians are conspicuously absent from the contest leaderboard. In this presentation, I will discuss the Netflix contest as well as the challenges and opportunities that it presents to the statistical community. In addition, I will describe some approaches to analyzing the Netflix data.

479 Compensating for Survey Nonresponse and Nonsampling Error ●

Section on Survey Research Methods

Wednesday, August 1, 2:00 pm–3:50 pm

An Assessment of the Effect of Misreporting of Phone Line Information on Key Weighted Estimates from Four National RDD Surveys

❖Ashley Bowers, University of Michigan, Survey Res Ctr, PO Box 1248 G317 Perry, Ann Arbor, MI 48187, *afbowers@umich.edu*; Jeffrey Gonzalez, Bureau of Labor Statistics

Key Words: Random-digit dialing (RDD), weighting adjustment, measurement error, sampling frame

Random-digit dialing (RDD) surveys typically incorporate a weighting factor that corrects for unequal probabilities of selection due to multiple eligible landlines per household. The information that is used for the phone line adjustment is self-reported in one or two questions at the end of the survey and the wording varies substantially across studies. Few studies have addressed whether these phone line estimates are subject to measurement error that might affect the weighting adjustment. We compare estimates from several ongoing RDD surveys to those based on extended series of phone lines items from five studies that are viewed as having higher quality estimates. We also develop a set of modified weights that account for misreporting of phone lines in four RDD surveys and then re-estimate key figures from each survey to determine if the measurement error impacts survey estimates.

Improvements in Coverage and Sample Weighting for Telephia's 2007 Attitude and Behavior Survey (TABS) of U.S. Cellular Phone Subscribers

❖Vicki Pineau, Telephia Inc., 101 Green St, San Francisco, CA 94065, *vpineau@telephia.com*; Jim Baer, Telephia Inc.; Christianne Pang, Telephia Inc.; Glenn Atias, Telephia Inc.; Elisa Chan, Knowledge Networks, Inc.

Key Words: Unacculturated Hispanics, Weight Trimming, Reduction in Sample Variance

The complex sample design of Tilapia's cornerstone survey on attitudes and behavior of the cell phone subscriber population called TABS requires substantial weighting. Additionally, post-stratification to reduce bias due to no coverage and nonsampling error further increases the sample variance for many outcomes of interest as compared to SRS. The observed level of variance in many survey outcomes however exceeded the control on the bias achieved and needed to be re-examined. Additionally, beginning January 2007, TABS will include a sample of the unacculturated Hispanic population that is not covered well (or at all) in most RDD and Internet samples. TABS will also include a sample of teens (13–17) as the teen cellular phone market has grown dramatically over the past few years in size and complexity of cell phone use.

Nonresponse in Surveys of the College Educated Revisited

❖Patrick Flanagan, U.S. Census Bureau, 1314 Bristol Ridge Place, Crownsville, MD 21032, *patrick.e.flanagan@census.gov*; Nancy Dexter, U.S. Census Bureau

Key Words: Nonresponse, Survey, logistic regression

The National Survey of College Graduates is one of three surveys that cover college and graduate degreed individuals in scientific fields. The Census Bureau has conducted one of these surveys, the National Survey of College Graduates (NSCG). In 1999, we conducted an analysis of nonresponse in the 1993 NSCG and compared the effect that various characteristics seem to have on nonresponse to results from other studies of the population in general. In this study, we repeated that analysis on the 2003 NSCG to see if the same effects were repeated and to take advantage of a wider array of frame variables. The results may be used to change the design of the nonresponse weight adjustment and possibly alter follow-up procedures.

Nonresponse in the Survey of Industrial Research and Development

❖Perez Adriana, The University of Texas at Houston Health Science Center, 80 Fort Brown SPH RAHC N102B, Brownsville, TX 78520, *adriana.perez@uth.tmc.edu*

Key Words: complex-survey, establishment survey, outcome, assessment, survey bias and precision

The Survey of Industrial Research and Development (SIRD) is important because provides statistics on research and development (R&D) by companies in the United States, which is used to evaluate the status of science and technology diffusion in the United States relative to other nations. Missing data is present in many of the SIRD key variables, varying from 1.1% to 68.2% (year 2003). Several imputation techniques (auxiliary ratio and auxiliary trend) are implemented to overcome missingness. Currently, estimates are reported for SIRD as if imputation values were observed values. A maximum likelihood estimate (MLE) of the data accounting for the uncertainty due to missing values is proposed. Empirical and simulations studies were developed to compare the results of current and proposed estimators through precision and accuracy measures (bias and mean square error).

Adjusting for Missing Covariates Using the EM Algorithm for Logistic Regression Models: A Study of the Effect of Childhood Overweight on the Age of Onset of Menses

❖Yi Liu, The Ohio State University, 484 Stinchcomb Dr APT 10, Department of Statistics, Columbus, OH 43202, *yliu@stat.ohio-state.edu*; Elizabeth A. Stasny, The Ohio State University; Pamela Salsberry, The Ohio State University; Patricia Reagan, The Ohio State University

Key Words: Discrete time hazard model, EM-algorithm

This paper investigated various approaches to correct for missing data in analyses that examine the influence of early weight states on age of onset of menses. A discrete time hazard model was used to analyze the dynamics of menses onset. The EM-algorithm was used to adjust for the missing values in one of the key covariates, the child's weight state at age 6/7 years.

Three different models were constructed and compared by fitting different structures on the covariate with missing values and on the missing indicator. Compared to the model based on complete data, the model adjusted for missingness predicted that girls were about 2 percentage points, or 6 percent, more likely to experience early onset menses (before age 11 years). Thus the empirical predictions of the model are sensitive to the treatment of missing data on childhood overweight as a risk factor for the timing of menses.

An Evaluation of Basic Calibration Estimators in Presence of Nonresponse

❖Yves Thibaudeau, U.S. Census Bureau, 1037 17th St S, Arlington, VA 22202, *yves.thibaudeau@census.gov*; Jun Shao, University of Wisconsin-Madison; Jeri Mulrow, National Science Foundation

Key Words: Calibration, nonresponse, Variance Estimation

Sarndal et al. (1992) discuss calibration estimators for complex surveys. These estimators rely on core survey information carried by covariates to reduce mean squared error. Sarndal and Lundstrom (2005) discuss using calibration estimators in presence of nonresponse. The Survey of Industrial Research and Development sponsored by the National Science Foundation exhibits several patterns of nonresponse involving its core items. But, some covariates are less often missing. We can take advantage of this situation and calibrate core items to reported information when available. Our ultimate objective is to evaluate jackknife variance estimation (Shao and Tu 1995) in presence of nonresponse for this type of calibration estimator.

The Nonparametric Regression with Missing Responses and Its Applications to the U.S. Census

❖Cheng Yong Tang, Iowa State University, 208A Snedecor Hall, Ames, IA 50014, *yongtang@iastate.edu*; Song Xi Chen, Iowa State University

Key Words: Nonparametric regression, missing values, imputation, discrete variables

The estimation of match rates in the U.S. census is a regression problem with missing responses and covariates which are mixtures of discrete and continuous variables. This paper considers the nonparametric regression in such problems. A nonparametric imputation approach for missing values are proposed which improves the estimation efficiency. Theoretical and simulation results demonstrated the performance of the proposed method. And some further applications of the nonparametric approach are discussed.

480 Methods and Analysis of Failure Time Data in Epidemiological Studies ●

Section on Statistics in Epidemiology

Wednesday, August 1, 2:00 pm–3:50 pm

Regression Analysis of a Disease Onset Distribution Using Diagnosis Data

❖Jessica G. Young, University of California, Berkeley, 711 Everett Street, El Cerrito, CA 94530, *jgerald@berkeley.edu*; Nicholas P. Jewell, University of California, Berkeley; Steven J. Samuels, University at Albany-SUNY

Key Words: Current status data, Proportional hazards, Uterine fibroids

We consider methods for estimating the effect of a covariate on disease onset when the data consist of right-censored data on diagnosis times and current status data on onset times for individuals not yet diagnosed. Dunson and Baird (2001) approached this problem using maximum likelihood, with the assumption that the ratio of the diagnosis and onset distributions is monotonic non-decreasing. We propose an alternative estimator that is computationally simpler and requires no assumptions on this ratio. A simulation study is performed comparing estimates from these two approaches, as well as that from a current status analysis ignoring diagnosis data. Results show that the Dunson and Baird estimator outperforms ours when the monotonicity assumption holds, but the reverse is true when this assumption fails. A data example is provided where the monotonicity assumption is seen to fail.

Semiparametric Methods for Estimating Cumulative Treatment Effects

❖Guanghui Wei, University of Michigan, Department of Biostatistics, Ann Arbor, MI 48109, *ghwei@umich.edu*; Douglas Schaubel, University of Michigan

Key Words: Confidence bands, Empirical processes, Integrated hazard, Stratification, Survival analysis, Time-dependent effect

Often in biomedical studies of time to an event, the treatment effect is not constant over time. In such cases, the cumulative treatment effect is usually of much greater importance than the instantaneous effect. We propose semiparametric methods for estimating the cumulative treatment effect in the presence of non-proportional hazards. The proposed estimator is based on the treatment-specific cumulative hazards estimated under a stratified Cox model. No functional form for the non-proportionality need be assumed. Asymptotic properties of the proposed estimators are derived and assessed through simulation. Simultaneous confidence bands are developed. The proposed method is applied to national organ failure and registry data. Extensions of the proposed methods to the recurrent event setting are considered.

Impact of Duration of Follow-Up Time on Predictors of Cardiovascular Disease in Survival Models

❖Lei Zhu, Boston University, 16 Walnut street Apt 42, Arlington, MA 02476, *leizhu@bu.edu*; Michael J. Pencina, Boston University; Ralph B. D'Agostino, Sr., Boston University

Key Words: cardiovascular disease, risk factors, proportional hazard model, duration of follow-up

Many cardiovascular disease (CVD) risk factors were identified over a specified time frame in the Framingham Heart Study. However, whether longer or shorter duration of follow-up affects the magnitude and significance of the effects remains uncertain. We used the Framingham Heart Study data to assess the impact of numerous predictors on the risk of CVD and potential changes of this impact during different durations of follow-up. Sex-specific multivariate Cox regression models were performed. We found that for some risk factors the magnitude or/and significance of the effect on CVD changed with time. For example, obesity in men was significantly related to CVD in a 5-year period, but lost significance in longer term. Our research suggests that risk assessment methods which assume constancy of effects regardless of duration of follow-up may overestimate or underestimate true relationships.

Semiparametric Regression with Time-Dependent Coefficient for Failure Time Data Analysis

❖Zhangsheng Yu, The Ohio State University, 320 W 10th Avenue Rm B108, Columbus, OH 48105, *zyu@sph.osu.edu*; Xihong Lin, Harvard School of Public Health

Key Words: Profile likelihood, Clustered survival data, Kernel smoothing, Marginal model

We propose a working independent profile likelihood method for the semiparametric time-dependent coefficient model with correlation. Kernel likelihood is used to estimate time-dependent coefficient. Profile likelihood for the parametric coefficient is formed by plugging in the nonparametric estimator. For independent data, the estimator is asymptotically normal and achieves the asymptotic semiparametric efficiency bound. We evaluate the performance of proposed nonparametric kernel estimator and the profile estimator, and apply the method to the western Kenya parasitemia data.

The Exclusion of Prevalent Cases from Incident Dementia Risk Factor Studies Sacrifices Lots of Information

❖ Daniel Tancredi, University of California, Davis, 2600 Kline CT, Davis, CA 95618, *dtancr1@yahoo.com*

Key Words: dementia epidemiology, case-2 interval censoring, parametric survival analysis, Fisher information, bias–variance tradeoff, risk factor estimation

Population-based studies of senile dementia risk factors typically collect and analyze data at two timepoints. Baseline data are used to study prevalent disease. Follow-up data are used to study incident disease among those free of disease at baseline. An alternative approach, using all of the data together in a single analysis, is tempting but rarely done because of concerns that the prevalent data are biased by selection effects (e.g., differential mortality). Could the risk factor analysis of incident disease benefit from including the prevalent cases? Parametric failure time regression models are specified for a simplified setting to address this question. Theory and computer simulations examine how truncation, interval censoring and selective mortality affect the bias and variance of risk factor estimators. Results show that including the prevalent cases offers big potential gains.

Comparison of Adjusted Survival Curves with Different Reference Populations

❖ Xu Zhang, Georgia State University, 30 Pryor Street 7th floor, Department of Mathematics and Statistics, Atlanta, GA 30309, *matxxz@langate.gsu.edu*

Key Words: Observational study, Selection bias, Cox model

The direct adjustment method is often used to construct the adjusted survival probabilities of different treatments, when imbalance of prognostic factors is present in the failure time data. The whole study cohort is the commonly selected reference population on which one applies direct adjustment. In this paper, we consider the problem of reference population selection. We apply the proposed procedures to real data and have the following conclusions. 1. In observational studies, the treatment groups may have distinct survival outcomes due to imbalance of prognostic factors. 2. When one wishes to compare the survival probabilities between two treatments, the comparison result is robust to various reference populations. 3. The group with inferior treatment can be used as the reference population to calculate the adjusted survival probabilities.

Proportional Odds Modeling of Discrete Survival Data with Application to Reproductive Study

❖ Huichao Chen, Emory University, 716 Carlyle Lake, Decatur, 30033, *hchen4@sph.emory.edu*; Limin Peng, Emory University; Amita K. Manatunga, Emory University; Michele Marcus, Emory University

Key Words: discrete survival, time to pregnancy, proportional odds, psuedo-likelihood

Using the proportional odds model, we develop a method to evaluate covariate effects on time-to-pregnancy (TTP), the length of time that it takes a couple to conceive. Specifically, to assess the effect of within-woman menstrual cycle variance on TTP, we first investigate the cycle variance via a log linear model, then implement the estimated variance as a covariate in modeling TTP. An unbiased estimating equation is used for modeling the variance in cycle length and a pseudo-likelihood procedure is used to estimate the parameters in the proportional odds model. It is shown that the resulting estimator is asymptotically normal, with a closed form variance-covariance matrix and can be consistently estimated by the usual empirical method. We illustrate the proposed approach using the data from the Mount Sinai Study of Women Office Workers (MSSWOW).

481 COPSS Awards and Fisher Lecture

The ASA, Committee of Presidents of Statistics Societies, IMS, ENAR, WNAR, SSC

Wednesday, August 1, 4:00 pm–5:50 pm

Presentation of Awards

❖ Madhuri S. Mulekar, University of South Alabama, 307 Univ Blvd ILB 325, Mathematics Statistics, Mobile, AL 36688, *mmulekar@jaguar1.usouthal.edu*

President's Award recognizes outstanding contributions to the profession of statistics by a young member of one of the participating societies. Sponsored and presented annually by COPSS. FN David Award recognizes a female statistician who exemplifies the contributions of Florence Nightingale David, an accomplished statistician. The award was established in 2001 and is presented biannually. Sponsored by COPSS and the Caucus for Women in Statistics. Snedecor Award honors an individual who was instrumental in the development of statistical theory in biometry which was published within three years of the date of the award. Dr. George W. Snedecor was instrumental in the world-wide use of statistical methods. It was established by COPSS in 1976 and awarded biannually. Fisher Lecturer was established by COPSS in 1963 to honor the contributions of Sir Ronald Fisher and the work of a present-day statistician. It recognizes the importance of statistical methods for scientific investigations and is awarded annually.

R. A. Fisher, Randomization, and Current Practice in Multicenter Clinical Trials

❖ Marvin Zelen, Harvard School of Public Health, 655 Huntington Ave, Boston, MA 02115, *zelen@hsph.harvard.edu*

Key Words: Fisher, Randomization, Clinical trials, Inference

R. A. Fisher's pioneering book, The Design of Experiments," published in 1935, laid the foundations of the modern design of experiments. The basic principles were randomization, replication, and local control. The importance of randomization, as a way of reducing biases in carrying out studies on humans, was early recognized by Bradford Hill. Today the randomized clinical trial is regarded as the "gold standard" for carrying out clinical trials. This lecture will discuss models in which the patient populations and hospitals are random samples or simply collections, and the implied limitations on making statistical inferences. In reality, the randomization process may serve as the only basis for making statistical inferences as the patients and hospitals are not random samples.

482 The Role of the ASA in MS Statistician Careers ✪

Committee on Membership Retention and Recruitment, ENAR, WNAR, Section on Survey Research Methods

Thursday, August 2, 8:30 am–10:20 am

The Evolution of an Engineer

❖Matthew Pocernich, National Center for Atmospheric Research, 3450 Mitchell Lane, Boulder, CO 80301, *pocernic@ucar.edu*

Key Words: career development, engineering, local chapter, involvement

Beginning my career as an environmental engineer who enjoyed statistics, the ASA helped me realize how much more there was to learn. This lead me back to graduate school for a degree in statistics. As a statistician I use the ASA as professional resource. Since ASA members are from many professions, I am exposed to research and methods not used in my specific field. While this is important, I now find myself using the ASA to find ways to use statistics in meaningful ways outside of work. This includes working with the local chapter in K–12 outreach programs as well as interest in the efforts in volunteerism.

Opportunities Offered by the ASA for a Statistician in a Management Role in the Pharmaceutical Industry

❖Jane B. Porter, Millennium Pharmaceuticals, Inc., 40 Landsdowne St, Cambridge, MA 02139, *porter@mpi.com*

Key Words: FDA, recuirting, pharmaceutical, management

The ASA offers a wealth of educational programs through Amstat News and newsletters from the local Chapters that can be shared with staff members. The FDA/Industry Workshop, in particular, always offers programs of interest to statisticians at all levels. In addition, the ASA offers opportunities for recruiting new statisticians. The presentation will focus on the benefit of ASA membership for pharmaceutical industry management and staff.

Transitioning into a Career as an MS Statistician: How the ASA Can Help

❖Monica Johnston, University of California, San Francisco, 185 Berry Street, Lobby 2, Suite 6600, San Francisco, CA 94107, *mjohnston@cc.ucsf.edu*

Key Words: transition, nontraditional, re-entry, salary, application, mid-career

At JSM 2006 I met people who were academically poised for a career as a MS statistician but were faced with common practical concerns: (1) which field of application to choose, (2) how to maintain a mid-career salary when taking their first job as a MS statistician, and (3) whether to continue their education or seek employment. This presentation will focus on the benefits of ASA membership for MS statisticians who are new to the field of statistics.

483 Advances in System Reliability

Section on Physical and Engineering Sciences, Section on Statisticians in Defense and National Security, Section on Risk Analysis, Section on Quality and Productivity

Thursday, August 2, 8:30 am–10:20 pm

Parametric and Nonparametric Flowgraph Models with Application to Bayesian Multistate Systems Reliability

❖Aparna V. Huzurbazar, University of New Mexico, MSC 03 2150 University of New Mexico, Dept of Math and Statistics, Albuquerque, NM 87131-0001, *aparna@stat.unm.edu*; Brian J. Williams, Los Alamos National Laboratory; David Collins, University of New Mexico

Key Words: recurrent event, time-to-event, covariates

Flowgraph models have been used for a variety of problems in multistate systems involving time-to-event data. This talk will focus on some recent extensions of these models. The first concerns modeling recurrent events and including covariates in traditional parametric flowgraph models. Recurrent event modeling is increasingly important in the area of component and system reliability both in terms of repair and replacement. Although the strength of flowgraph models is that they model observable waiting times, allow for a variety of different distributions to be used within a stochastic network, and allow for all orders of loops, their weakness is that to date they do not allow for continuous time covariates. The second concerns extending these models to use nonparametric methods.

Integrating System Engineering Methods into Reliability Estimates

❖Andrew Koehler, Los Alamos National Laboratory, 1820 Hopi Road, Santa Fe, NM 87505, *akoehler@lanl.gov*

Key Words: system engineering, reliability, simulation, lifecycle design

Traditionally, system engineering and system reliability estimation have been separate activities. Reliability methods have focused on predicting how a particular inventory of systems will age and for how long a given percent will remain functional. This effort has been predicated on the assumption that a system architecture is defined, and that a sufficient body of testing (whether derived from functional or component sources) exists. In contrast, system engineering has focused on the determination of an architecture that can be validated against customer requirements. In this paper, I will discuss these issues, and describe how concurrent system engineering approaches can be included in reliability prediction efforts.

Resource Allocation Strategies for Reliability of Complex Systems

❖Christine Anderson-Cook, Los Alamos National Laboratory, P.O. Box 1663 MS F600, Los Alamos, NM 87545, *c-and-cook@lanl.gov*

Key Words: system modeling, sampling, design of experiments, bayesian analysis

When modeling the reliability of complex systems, many different data sources, from full-system tests to component quality assurance measures to accelerated test data, can be utilized to improve system and component reliability estimates. While much of the previous work has focused on developing a working model for existing data, this talk considers what data collection strategies would be most advantageous for optimizing system reliability precision in the future. A metric to quantify good performance of future analyses is specified. Using this metric, the strategies for determining which data should be collected are dependent on current data available, the reliability of individual components and the relative cost of new data sources. A simple example will illustrate the approach.

484 Statistical Validation of Surrogate Endpoints ● ✪

Biopharmaceutical Section, Biometrics Section, Committee on Applied Statisticians, WNAR

Thursday, August 2, 8:30 am–10:20 am

Unified Approaches for Surrogate Marker Evaluation from Multiple Randomized Clinical Trials

❖Geert Molenberghs, Hasselt University, Center for Statistics, Agoralaan 1, Diepenbeek, B3590 Belgium, *geert.molenberghs@uhasselt.be*

Key Words: surrogate markers, biomarkers, meta-analysis, surrogate endpoints, information theory

The validation of surrogate endpoints has been initially studied by Prentice and Freedman. Noting operational difficulties, Buyse and Molenberghs proposed instead to use jointly the within-treatment partial association of true and surrogate responses, and the treatment effect on the surrogate relative to that on the true outcome. In a multicenter setting, these quantities can be generalized to individual-level and trial-level measures of surrogacy. Buyse and colleagues have proposed a meta-analytic framework to study surrogacy at the trial and individual-patient levels. Variations for various endpoints have been developed. Efforts have been made to converge to a common framework. This includes a so-called variance reducation factor and an information-theoretic approach. Work has been done regarding sample size assessment, leading to the surrogate threshold effect.

Evaluating and Using Surrogate Endpoints in Clinical Trials: A Bayesian Approach

❖Mary K. Cowles, The University of Iowa, 241 Schaeffer Hall, Department of Statistics and Actuarial Science, Iowa City, IA 52242-1409, *kate-cowles@uiowa.edu*; Qian Shi, The University of Iowa

Key Words: accelerated failure time model, censored data, proportional hazards model, Wishart distribution

Surrogate endpoints in clinical trials are biological markers or events that may be observed earlier than the clinical endpoints (such as death) that are actually of primary interest. We propose Bayesian models for use in settings in which the true endpoint of interest is time to a clinical event and there are one or more continuous-valued markers that may be useful surrogates, either alone or in combination. The time-to-event component of our models may be either parametric or semi-parametric. Our parametric models can produce posterior predictive distributions for the event times of individuals with censored data, given values of markers and other covariates.

Surrogate Endpoint Validation: Statistical Elegance versus Clinical Relevance

❖Daniel J. Sargent, Mayo Clinic, 200 1st St SW, Rochester, MN 55901, *sargent.daniel@mayo.edu*

Key Words: Surrogate Endpoint, Meta Analysis

A variety of approaches have been proposed to provide formal and informal validation of proposed surrogate markers. To achieve true clinical impact, the validation must convince statistical, regulatory, and clinical communities. In this talk I argue, through an actual data example, that the best approach is not a single method but a multi-faceted exploration. Effective validation will use at least two types of approaches (1) those that directly appeal to clinicians but with less statistical foundation (focusing on graphics and simple data summaries), and (2) those arising from statistical modeling (such as multivariate models), but which are more difficult to interpret clinically. Both methods are maximally powerful when using the meta-analytical approach.

485 Statistics and Finance ● ✪

IMS

Thursday, August 2, 8:30 am–10:20 am

Large Dimensional Covariance Matrix Estimation for Asset Pricing and Risk Management Using a Factor Model

❖Jianqing Fan, Princeton University, Department of Operations Res and Fin Eng, Princeton, NJ 08544, *jqfan@princeton.edu*; Yingying Fan, Princeton University; Jinchi Lv, Princeton University

Key Words: Covariance estimation, Factor Model, High-dimensionality, Portifolio Allocation, Risk Management

Large dimensionality comparable to the sample size is a common feature as in modern portfolio allocation and risk management. In this paper we examine the covariance matrix estimation in the asymptotic framework that the dimensionality p grows with sample size. Motivated by financial economics theory, we propose to use a multi-factor model to reduce the dimensionality and to estimate the covariance matrix among those assets. Under some basic assumptions, we have established the rate of convergence and asymptotic normality for the proposed covariance matrix estimator. The performance is compared with the sample covariance matrix. We identify the situations under which the factor approach can gain substantially the performance and the cases where the gains are only marginal.

Do Jumps Matter?

❖Lan Zhang, University of Illinois at Chicago, 60607, *lanzhang@uic.edu*

Key Words: high frequency data, jump

The detection of jumps is a major focus of current work on high-frequency financial data. The present paper argues that the distinction between jumps and continuity may only be moderately useful. It is, both practically and theoretically, better to consider small jumps as part of the continuous component of the process, and only to separate out the big jumps. A mathematical construction is given, and estimation methods are discussed for the case when there is no microstructure noise.

Testing for Jumps in a Discretely Observed Process

❖Yacine Ait-Sahalia, Princeton University, Department of Economics, 26 Prospect Ave, Princeton, NJ 08540, *yacine@princeton.edu*; Jean Jacod, Universite Pierre and Marie Curie

Key Words: Jumps, Diffusions, High frequency, Semimartingales, Discrete sampling

We propose a new test to determine whether jumps are present in asset returns or other discreetly sampled processes. As the sampling interval tends to 0, our test statistic converges to 1 if there are jumps, and to another deterministic and known value (such as 2) if there are no jumps. The test is valid for all Ito semimartingales, depends neither on the law of the process nor on the coefficients of the equation which it solves, does not require a preliminary estimation of these coefficients, and when there are jumps the test is applicable whether jumps have finite or infinite activity and for an arbitrary Blumenthal-Getoor index. We finally implement the test on simulations and asset returns data.

486 Practical Bayesian Nonparametric Methods ● ✪

Section on Bayesian Statistical Science, Section on Nonparametric Statistics

Thursday, August 2, 8:30 am–10:20 am

Bayesian Nonparametric Modeling for Well-Calibrated Location and Scale Inference with Skewed and Heavy-Tailed Data

❖ David Draper, University of California, Santa Cruz, Department of Applied Mathematics and Statistics, Baskin School of Engineering, 1156 High Street, Santa Cruz, CA 95064, *draper@ams.ucsc.edu*

Key Words: bootstrap, Dirichlet-process, mixture modeling

The bootstrap is a popular frequentist nonparametric technique for creating interval estimates, which claims to produce well-calibrated intervals no matter what the underlying distribution F is. However, since it's based solely on the empirical cumulative distribution function, it has no information about F beyond the largest data value Y(n). When n is moderate and F is heavy-tailed and/or heavily skewed, this can ignore much of the "weight" of the underlying distribution, leading to (extremely) poor calibration for location and scale functionals. In this talk I will describe the use of Bayesian nonparametric (Dirichlet process mixture) modeling to produce well-calibrated location and scale intervals, even when n is quite small and the (unknown) data-generating F is quite skewed and/or heavy-tailed.

Nonparametric Modeling for Spatial Functional Data Analysis

❖ Alan E. Gelfand, Duke University, ISDS, Durham, NC 27708-0251, *alan@stat.duke.edu*

Key Words: Dirichlet processes, hierarchical model, random functions

Recently there has been increased interest in flexible modeling for functional data analysis. We consider a strategy based upon Dirichlet process mixing. But then, we may encounter unknown functions at various spatial locations. However, we might expect that functions closer to each other in space will be more similar than those farther apart. We would like to develop a spatial process model that is again, nonparametric but captures this behavior. Here we can employ spatial Dirichlet processes. Altogether, we create a Bayesian nonparametric model for spatial functional data analysis. We illustrate the application of this modeling methodology with a dataset involving temperature vs. depth relationships at many locations in the Atlantic Ocean.

Bayesian Semiparametric Modeling Based on Mixtures of Polya Trees

❖ Timothy E. Hanson, The University of Minnesota, Div. of Biostatistics, A460 Mayo Building, MMC 303, School of Public Health, 420 Delaware St SE, Minneapolis, MN 55455, *hanson@biostat.umn.edu*

Key Words: Nonparametric, mixed model

Relative to Dirichlet process mixtures (DPM) and other nonparametric prior processes, the Polya tree and mixtures of Polya tree (MPT) priors have seen relatively limited use. In this talk I will briefly provide some background on MPTs and implementation notes, then move on to a number of real data examples where the MPT model improves prediction relative to a parametric model or equivalent DPM model. Examples will include survival models with time dependent longitudinal trajectories, survival models with spatial frailties, diagnostic serology test data modeled subject to a stochastic order constraint, dependent Polya tree processes, and Polya tree mixtures for meta-analysis in non-normal settings.

Modeling with Infinitely Divisible Moving Averages

❖ Robert L. Wolpert, Duke University, 211c Old Chem Building, Statistical Science, Durham, NC 27708-0251, *wolpert@stat.duke.edu*

Key Words: LÈvy random fields, alpha stable processes

Continuous-parameter moving averages of independent increment processes are necessarily infinitely divisible—often Gaussian, but more recently moving averages of stable, gamma, and other ID processes have proven useful in a wide variety of applications, not only in one dimension but also for spatial and higher-dimensional settings. Basic background, computational methods, and applications of ID (and in particular LÈvy) random fields are presented as prior distributions for nonparametric Bayesian analyses.

487 Profile Monitoring ● ✪

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Thursday, August 2, 8:30 am–10:20 pm

Profiles: Monitoring the Variance Components

❖ Shilpa Gupta, Arizona State University, Department of Industrial Engineering, Tempe, AZ 85287-5906, *shilpa.gupta@asu.edu*; Douglas C. Montgomery, Arizona State University

Key Words: Profile Monitoring, Residual Control Charts, CUSUM

The key challenge in profile monitoring is to correctly account for the correlation within profiles. Most of the methods in the literature provide optimal monitoring schemes ignoring the correlation. A residual control chart method is proposed to and illustrated with examples.

Monitoring Correlation Within Nonlinear Profiles Using Mixed Models

❖ Willis Jensen, W.L. Gore & Associates, Inc., 3750 West Kiltie Lane, Flagstaff, AZ 86003-2400, *wjensen@wlgore.com*; Jeffrey B. Birch, Virginia Polytechnic Institute and State University

Key Words: Multivariate Statistical Process Control, Nonlinear Model, Phase I, Blups, T-squared Statistic

Profile monitoring is a technique in quality control best used where process data follow a profile at each time period. Previous work on monitoring of nonlinear profiles has assumed that the measurements within a profile are uncorrelated. To relax this restriction we propose the use of a nonlinear mixed model (NLMM) to monitor the nonlinear profiles to account for the correlation structure. We evaluate the effectiveness of fitting separate nonlinear regression models to each profile in Phase I control chart applications for data with uncorrelated errors and no random effects. We make the same evaluation for data with random effects. Our proposed approach uses the separate nonlinear regression model fits to obtain the NLMM fit. This NLMM approach results in charts with good abilities to detect changes in Phase I data and has a simple to calculate control limit.

A General Framework for Monitoring of Dynamic Systems

❖Kwok-Leung Tsui, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, *ktsui@isye.gatech.edu*; Wei Jiang, Stevens Institute of Technology; Tom Au, AT&T Labs - Research

Key Words: activity monitoring, data mining, dynamic systems, profile monitoring, complex data

This talk discusses strategies and techniques in data modeling, forecasting, and activity monitoring. Activity monitoring refers to detection of interesting events that require actions (e.g., detection of customer churn, credit card or insurance fraud, and computer intrusion). We will propose a general strategy for modeling, forecasting, and monitoring of dynamic systems. In particular, we will discuss a statistical process control approach for business activity monitoring. We will also discuss a churn detection procedure for customer profile modeling. Several examples and case studies in telecom and service industries will be used to illustrate the proposed methods.

488 Innovative Designs and Analysis in Clinical Trials: Fully Harnessing Patient Information Accrued During a Trial ● ✪

Section on Government Statistics, ENAR, Biometrics Section, Biopharmaceutical Section, WNAR

Thursday, August 2, 8:30 am–10:20 am

Adaptive Dose-Finding Designs for Single and Dual Agent Combinations Using Accumulated Data from All Patients

❖Sumithra J. Mandrekar, Mayo Clinic, Rochester, MN 55901, *mandrekar.sumithra@mayo.edu*; Daniel J. Sargent, Mayo Clinic

Key Words: Phase I, Trinomial Outcome, CR Model, Efficacy, Toxicity, CRM

Historically, designs for dose seeking trials have been geared towards finding the maximum tolerated dose, with safety as the primary outcome. Agents with specific biologic targets with unknown dose efficacy curves and limited toxicity profile require novel designs. We present designs for identifying an optimal dose or combination region utilizing both toxicity and efficacy data. A continual reassessment method with straightforward dose selection criteria using accumulated data from all patients treated thus far is employed based on the flexible continuation-ratio model. A generalization of the single agent model allowing for separate toxicity and efficacy curves for each agent to generate a dose outcome surface for the two agent combination is used. Our simulation studies demonstrated that the proposed designs have favorable operating characteristics under a variety of scenarios.

Adaptive Design and Analysis of Clinical Trials

❖Yu Shen, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 447, Department of Biostatistics, Houston, TX 77030, *yshen@mdanderson.org*; Yi Cheng, Indiana University South Bend

Key Words: Clinical trial, adaptive, estimation, inference

In this talk, we will describe statistical methodology in clinical trial designs that are adaptive, allowing flexibility for sample size determination, yet maintaining scientific validity. The methods will be applicable to various areas of health research, especially cancer clinical trials. Adaptive clinical trials continually update the design using cumulative information. Two types of adaptive sequential designs will be discussed within the framework of Bayesian and frequentist methods. The proposed methods can be used when outcomes are continuous, binary and time-to-event variables. Adaptive trials terminated early may exaggerate treatment benefits or harm in sequential designs. The investigation of bias adjusted estimators for various adaptive designs will be presented too.

Progress in Using Adaptive Designs for Clinical Trials

❖Donald A. Berry, The University of Texas M.D. Anderson Cancer Center, Department of Biostatistics, 1515 Holcombe Boulevard, Unit 447, Houston, TX 77030-4009, *dberry@mdanderson.org*

Key Words: Adaptive design, Clinical trials, Innovative designs, Progress, Regulations, Successes and failures

I will recount successes in developing innovative designs in clinical trials, and successes and failures in running them. I will distinguish between trials in phases I and II and pivotal phase III trials, especially as regards regulatory concerns. I will address the future of statistical ideas in clinical research quite generally, with special reference to adaptive design.

489 Medication Compliance: New Methods and Applications ● ✪

ENAR, WNAR, Section on Health Policy Statistics, Biometrics Section

Thursday, August 2, 8:30 am–10:20 am

Noncompliance with Drug-Dosing Regimens: An Underestimated Bias in Drug Trials and Clinical Practice

❖Bernard Vrijens, Pharmionic Systems Ltd., Rue des Cyclistes FrontiÈre 24, VisÈ, BE-4600 Belgium, *bernard.vrijens@pharmionic.com*

Key Words: Adherence, Compliance, Persistence, Electronic Monitoring

The Pharmionic Knowledge Centre is a data warehouse on more than 20000 ambulatory patients whose dosing histories during studies of varying lengths have been electronically compiled. Electronic Medication Event Monitors were used to record the times and dates of drug dosing during the course of >90 drug trials between 1990 and 2006. Chapter headings in the British National Formulary served to categorize fields of treatment. Study durations ranged from 30 to 1400 days. Patterns of deviation from prescribed dosing regimens varied widely, but were almost entirely markedly skewed toward longer dosing intervals than prescribed, i.e. under-dosing, in every field of treatment. Drug holidays, and early cessation of dosing are common features in ambulatory patients, and likely are frequent sources of low response and high variability in response to the prescribed dosing regimen.

Dose-Timing Intervals of Antiretroviral Medications and Virologic Responses

❖Honghu Liu, University of California, Los Angeles, Division of General Internal Medicine & Health Services Research, Department of Medicine, School of Medicine, Los Angeles, CA , *hhliu@mednet.ucla.edu*; Loren G. Miller, University of California, Los Angeles; Ron D. Hayes, University of California, Los Angeles; Carol E. Golin, The University of North Carolina at Chapel Hill; Junfeng Sun, University of Nebraska Medical Center; Tongtong Wu, University of California, Los Angeles; Neil S. Wenger, University of California, Los Angeles

Key Words: Adherence, Dose timing, Genotypic sensitivity, Viral load, HIV/AIDS

Adherence is a main predictor of the effectiveness of antiretroviral medications in HIV treatment. But studies have focused primarily on the percent of prescribed doses taken (percent adherence). We collected detailed adherence data including dose timing and patients' virologic responses. For 48 weeks, adherence and virologic outcomes were collected every 4 weeks, and demographics and other measures were collected at baseline and weeks 8, 24, and 48. We constructed 8 dose timing error (DTE) measures and evaluated their associations with virologic outcomes. Repeated measures mixed effect models were fitted to evaluate the predicting power of these DTE measures. Among 52,036 electronically recorded doses from 122 patients, DTE measures significantly predicted virologic outcomes. Six DTE measures were significantly predictive of virologic outcomes after controlling for percent adherence.

Analyzing the Effects of Compliance Measured with Error in Randomized Trials

❖ Els Goetghebeur, Gent University, Krijgslaan 281 S9, Gent, International 9000 Belgium, *els.goetghebeur@ugent.be*; Stijn Vansteelandt, Ghent University; Manoochehr Babanezhad, Ghent University

Key Words: causal inference, compliance, measurement error, structural models

For active drugs, exposure patterns are on the causal pathway from prescription to response, and tend to vary. Structural (generalized) mean models allow for randomization based estimates of their causal effects in terms of potential outcomes. Their estimating equations remain unbiased with mean zero errors on compliance measures, but systematic errors are common. With known (covariate specific) mean (and variance) for the error on compliance, we can build (better) adjusted unbiased structural linear and approximate logistic mean estimating equations. We also propose a new estimator relying on an additional instrumental variable to estimate the effect of true exposure on the response of interest. The unbiased estimating equations are obtained in the structural linear mean context and modified to obtain better mean squared errors. We explore their value in several randomized clinicals.

490 Advances in Statistical Aggregation ✪

IMS, International Indian Statistical Association

Thursday, August 2, 8:30 am–10:20 am

Aggregation Using Empirical Risk

❖ Sara A. van de Geer, ETH Zurich, Seminar fuer Statistik, LEO D11, Zurich, 8092 Switzerland, *geer@stat.math.ethz.ch*

Key Words: empirical risk, excess risk, oracle inequality

We study the situation where estimators built on a training set are combined using the empirical risk in a test set. In model selection, we examine when in the selected estimator has excess risk close to the optimal one, and to what extend this depends on a bound for (moments of) the initial estimators, and on the margin behavior (the latter being in particular of interest in classification problems). In linear or convex aggregation, the excess risk of the aggregated estimator is close to the optimal excess risk if aggregation is carried out over a limited number of initial estimators. When this number is very large however, a regularized empirical risk should be used. We study the Lasso penalty and prove oracle inequalities for general, possibly unbounded, loss functions.

Sparsity Bounds in Empirical Risk Minimization

❖ Vladimir Koltchinskii, Georgia Institute of Technology, School of Mathematics, Atlanta, GA 30332-0160, *vlad@math.gatech.edu*

Key Words: sparsity, empirical risk minimization, ℓ_p-penalization, excess risk, oracle inequalities

A number of problems in Statistical Learning Theory, such as regression and pattern classification, can be formulated as penalized empirical risk minimization over a linear span of a very large dictionary of functions. The complexity penalties are often based on ℓ_p-norm of the vector of coefficients, most often, with $p=1.$ The central problem that will be discussed in this talk is to show that if the true solution of the problem is "approximately sparse," then its empirical solution is also "approximately sparse." Several "sparsity bounds" that provide mathematical description of this phenomenon will be considered and the influence of the sparsity on excess risk bounds and oracle inequalities will be also discussed.

Aggregating Algorithms for Individual Sequences

❖ Nicoló Cesa-Bianchi, Universitá degli Studi di Milano, Dipartimento di Scienze dell'Informazione, via Comelico 39, Milano, 20135 Italy, *cesa-bianchi@dsi.unimi.it*

Prediction with expert advice is concerned with the design and analysis of aggregating algorithms without imposing any probabilistic assumption on the data-generating mechanism. The performance bounds that can be proven in such a general framework are surprisingly strong. Typically, these bounds depend on quantities that are empirical versions of analogous entities arising in statistical aggregation problems. We will review a number of results in prediction of individual sequences, including some recent results where the regret is bounded in terms of empirical variances. We will also describe some simple techniques for converting any individual sequence aggregation bound into a corresponding bound for i.i.d. sources. These techniques are effective, in the sense that little processing is required besides the computation performed by the original individual sequence algorithm.

Sparsity Oracle Inequalities for the Lasso

❖ Marten Wegkamp, Florida State University, Tallahassee, FL 32306-4330, *wegkamp@stat.fsu.edu*; Alexandre Tsybakov, Paris VI; Florentina Bunea, Florida State University

Key Words: Lasso, Sparsity, Mutual coherence, Oracle inequalities, Empirical risk minimization

We study oracle properties of the lasso estimator in a nonparametric regression setting with random design. Our goal is to estimate the regression function via linear combinations of elements in some dictionary and to estimate its performance in terms of the L2 risk. An oracle that knows the regression function would be able to tell us in advance the sparsest approximating submodel with the smallest risk. We show that the lasso estimator behaves like the oracle. The results are valid even when the dimension of the initial model is (much) larger than the sample size and the regression matrix is not positive definite. Instead we introduce a new concept of local mutual coherence. Our results can be applied to high-dimensional linear regression, to nonparametric adaptive regression estimation and to the problem of aggregation of arbitrary estimators.

491 Statistical Methods and Applications in Genomics ● ✪

IMS, Section on Statistics in Epidemiology, Section on Health Policy Statistics, Biometrics Section, Biopharmaceutical Section, WNAR

Thursday, August 2, 8:30 am–10:20 am

PfCluster: A New Profile Cluster Analysis Procedure

❖ Jiayang Sun, Case Western Reserve University, Department of Statistics, 10900 Euclid Ave, Cleveland, OH 44106-7054, *jsun@case.edu*; Yaomin Xu, The Cleveland Clinic

Key Words: Profile cluster analysis, coherence index, pathway, regulatory network, gene expression, P-value

Properly clustering gene profiles collected under different conditions or from a longitudinal study is important in building gene regulatory networks and for finding gene pathway information for certain diseases. In this paper, we propose a new clustering technique called PfCluster. PfCluster can uncover clusters using a distance from a class of biologically meaningful distances. "Within" and "Between" coherence indices are developed to measure how coherent resulting clusters are. The null distributions of coherence indices are given. These indices provide some measure of the integrity of resulting groups, and an inferential procedure for deciding where to cut a dendrogram. An efficient filtering method is also provided to filter out the idle genes that have a flat profile pattern. Applications to real data and some simulation results will also be presented.

A Simple and Improved Correction for Population Stratification in Genetic Association Studies

❖ Michael Epstein, Emory University, School of Medicine, 615 Michael Street Suite 301, Atlanta, GA 30322, *mepstein@genetics.emory.edu*

Case-control studies of disease-gene association must often correct for potential bias resulting from the effects of population stratification, which originates from the coupling of allele-frequency heterogeneity to disease-risk heterogeneity within a population. Such stratification remains an important issue in association studies, even within populations thought to be homogeneous in nature. To provide an improved correction for stratification, we propose a simple two-step procedure. First, we construct a model for the odds of disease given a set of marker loci (excluding the loci of interest) that can provide information on population substructure. We use the first-step model to assign subjects to strata based on their odds of disease given marker genotypes. Then, in the second step, we test for disease-gene association adjusting for the strata using logistic regression.

Prior Elicitation and Variable Selection in Regression Models with High-Dimensional Data

❖ Joseph G. Ibrahim, The University of North Carolina at Chapel Hill, Department of Biostatistics, Chapel Hill, NC 27614, *ibrahim@bios.unc.edu*; Mayetri Gupta, The University of North Carolina at Chapel Hill

Key Words: Generalized Linear Models, Survival Analysis, g-prior, Microarray Data

One of the most important modern day challenges in analyzing high dimensional data (such as microarray data) jointly with continuous, discrete, or time-to-event data is that one immediately encounters the $p > n$ problem, in which the number of covariates in the regression model greatly exceeds the number of subjects. In this talk, we develop a methodology for the specification of a class of prior distributions for generalized linear models and survival models that accommodate the $p > n$ paradigm. The class of proper prior distributions are based on the generalization of the g-prior as well as having a 'ridge parameter' that facilitates propriety for $p > n$. The resulting prior has the flavor and operating characteristics of ridge regression. Various properties of the prior are discussed and a real dataset is analyzed to illustrate the proposed methodology.

492 Network Science

Section on Statistical Computing, Section on Bayesian Statistical Science

Thursday, August 2, 8:30 am–10:20 am

Bayesian Inference for a Longitudinal Social Network Model

❖ Mark S. Handcock, University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195-4322, *handcock@stat.washington.edu*

Key Words: random graph models, stochastic process, MCMC

In studies of social networks recent emphasis has been placed on random graph models where the nodes represent individual social actors and the edges represent a specified relationship between the actors. Much progress has been made on exponential family models for cross-sectional networks, and some has been made on related models for networks observed longitudinally. A fundamental goal of social network theory is to represent the processes of network formation over time. The theory of balance developed by Heider posits that networks evolve towards structural balance. An alternative theory due to Simmel has the triad as the basic social unit. We develop Bayesian inference for a model that represents the level and dynamics of Heiderian balance and Simmelian stability, and access the empirical support for the two theories. This is joint work with David Krackhardt and Martina Morris.

Network Survivability

❖ Nozer D. Singpurwalla, The George Washington University, Department of Statistics, 2140 penn. Av., Washington, DC 22207, *nozer@gwu.edu*

Key Words: Interdependence, Interaction, Social Networks, Graph Theory

Networks, be they social, communications, or transportation, can experience failures. Such failures can be due to adversarial or natural actions. Assessing the survivability of networks is therefore an activity that is germane. Graph and network theorists have assessed network reliability by assuming that the failure propensities of the network's nodes are known, and that the failures are independent. This is an idealization. In this talk, I shall make a distinction between network survivability and network reliability and then point out issues that need to be addressed to obtain the former. I shall make the argument that node interdependence is hierarchical and that independence can only be claimed under independence at both the two stages of the hierarchy. The role of Markov chain Monte Carlo for putting my ideas to work so computable results can be obtained and highlighted.

A Review of Statistical Models for Networks with a Focus on Imputation

❖ Stanley Wasserman, Indiana University, Ballantine Hall 744, Bloomington, IL 47405, *stanwass@indiana.edu*; Douglas Steinley, University of Missouri-Columbia; Garry Robins, University of Melbourne

Key Words: Network Science, Categorical Data, Imputation

This talk highlights the wide range of statistical analyses that are part of network science. Of particular importance are the exponential family of

random graph distributions, known as p^*, and recent work on robustness and resistance of network data when actors and/or relational ties are missing or removed.

493 Consulting from a Distance: Web-Based Analytics for Statistical Consulting ● ✪

Section on Statistical Consulting, WNAR

Thursday, August 2, 8:30 am–10:20 am

Web-Based Analytics in Marketing

❖Samaradasa Weerahandi, Time Warner, 135 West 50th St, 10207H, New York, NY 10020, *weerahandi@aol.com*

Key Words: Analytics, Business Intelligence, Splus Server, SAS Server, Web-based Analysis

Most statistical applications, ranging from marketing to pharmaceutical applications, involve pulling data from one or more databases, cleaning data, and analyzing them. Performing most of these tasks can be accomplished by developing web analytics that automatically pull data and performing such analyses before delivering them to the user. With substantially reduced time to deliver reports, web analytics can highly contribute to Statistical Consulting, especially in recurrent analyses, and they can maximize the productivity and efficiency of many organizations requiring data based decisionmaking. Web publishing of Statistical Analytics is easily accomplished via commercially available application servers such as the SPlus Server and the SAS Server. After a discussion of above issues, this presentation will provide a demo of some Marketing Analytics.

Use of Remote Consulting in a University Environment

❖Edward D. Rothman, University of Michigan, 3560 Rackham, Ann Arbor, MI 48109, *erothman@umich.edu*

Key Words: remote consulting, education, software

The Center for Statistical Consultation at the University of Michigan faced with a shortage of space and an increased demand for consulting help uses a remote consulting package, CENTRA, to help. With unknown demand, we began to roll this option out. The package allows access to computers in both our location and the client's, white board, voice, and a facial image. Constraints had to be imposed on the clients who would want access because 50% of our clients are graduate students, and researchers would prefer access that does not involve parking! The limitations of the interaction and the advantages will be described. Cost is an issue but with an institution with shared access and a teaching option many others have come together to make this a possibility. Other software options will also be discussed.

Remote Consulting on Linear Mixed Models at a University Consulting Center

❖Kathleen Welch, University of Michigan, CSCAR 3050 Rackham Bldg, 915 E Washington St, Ann Arbor, MI 48109-1070, *kwelch@umich.edu*; Brady West, University of Michigan

Key Words: remote consulting, linear mixed models, statistical consulting

The Center for Statistical Consultation and Research at the University of Michigan (CSCAR) has many off-campus clients. We have implemented a flexible remote consulting model using CENTRA software, allowing us to consult on statistical and data-base management issues for clients located anywhere with an internet connection. Fitting Linear Mixed Models (LMMs) for longitudinal or clustered data presents unique challenges in a remote setting. Having a "hands-on" connection with a client allows the consultant to work with data, provide guidance in model setup, interpret results, and work with diagnostics. The client can see and use different statistical packages on the consultant's desktop. Notes may be kept on the client's computer to document the session. Problems with LMM model convergence may be more readily diagnosed in a remote consulting session than via email or over the phone.

How to Move from Desktop to Web Analytics

❖Dilip Patel, Time Warner, 135 W 50th Street, New York, NY 10020, *dilip_patel@timeinc.com*

Key Words: Web Analytic, Business Intelligence, Desktop analytic, database, splus server, sas server

In today's competitive business world one has to move from traditional desktop based analytical reporting to web based or central report distribution mode. To achieve this one has to go through various steps of building a central Business Intelligence Portal. In this session I will drive the audience through some steps we took at Time Inc to build a web based business portal covering simple summary reports to advance analytical report. A simple example of how to migrate a desktop based code (e.g. Splus) to a Web enabled analytic using an application server (e.g. Splus Server) will be presented.

494 Modeling Data from Brain Imaging Studies ● ✪

Biometrics Section, WNAR

Thursday, August 2, 8:30 am–10:20 am

Dynamic Positron Emission Tomography Data-Driven Analysis Using Sparse Bayesian Learning

❖John Aston, Academia Sinica, Institute of Statistical Science, 128 Academia Road Sec 2, Taipei, 115 Taiwan, *jaston@stat.sinica.edu.tw*; Jyh-Ying Peng, Academia Sinica

Key Words: Sparse Bayesian Learning, PET, Exponential Bases

A method is presented for the analysis of dynamic PET data using sparse Bayesian learning. Parameters are estimated in a compartmental framework using an over-complete exponential basis set and sparse Bayesian Learning. The technique is applicable to analyses requiring either a plasma or reference tissue input function and produces estimates of the system's macro-parameters and model order. The method is applied to the estimation of parametric images of neuroreceptor radioligand studies.

FMRI Statistical Brain Activation from k-Space Data

❖Daniel Rowe, Medical College of Wisconsin, Department of Biophysics, 8701 Watertown Plank Road, Milwaukee, WI 53226, *dbrowe@mcw.edu*

Key Words: complex-valued, k-space, fMRI, brain activation

In fMRI brain activation is commonly performed in terms of voxel time series measurements after image reconstruction. The image reconstruction and statistical activation processes are treated separately. The relationship between complex-valued k-space measurements and complex-valued image measurements is summarized. The voxel time-series measurements are written in terms of spatio-temporal k-space measurements. FMRI activation is determined in image space in terms of the original k-space measurements. Additionally, the spatio-temporal covariance between reconstructed complex-valued voxel time series can be written in terms of the spatio-temporal covariance between complex-valued k-space measurements. This allows one to utilize the originally measured data in its more naturally acquired state rather than in a transformed state. This allows statistical modeling in k-space.

State-Space Models for Optical Imaging

❖Kary Myers, Los Alamos National Laboratory, 2273 47th St Apt B, Los Alamos, NM 87544, *karymyers@gmail.com*; Anthony E. Brockwell, Carnegie Mellon University; William F. Eddy, Carnegie Mellon University

Key Words: optical imaging, linear state-space models, Kalman filtering

Measurement of stimulus-induced changes in activity in the brain is critical to the advancement of neuroscience. These measurements are usually corrupted by artifacts, complicating interpretation of the signal. In the context of optical imaging of intrinsic signals, two primary sources of corruption are the heartbeat and respiration cycles. We introduce a new linear state-space framework which uses the Kalman filter to remove these artifacts from optical imaging data. The method relies on a likelihood-based analysis under the specification of a formal statistical model, and allows for corrections to the signal based on auxiliary measurements of quantities closely related to the sources of contamination, such as physiological processes. Working with data collected by our collaborators, we demonstrate the method on data collected in an optical imaging study of a cat's brain.

Assessing Homogeneity in Brain Regions Using Imaging Data

❖R. Todd Ogden, Columbia University, 722 W 168th St, Department of Biostatistics, 6th floor, New York, NY 10032, *to166@columbia.edu*; Huiping Jiang, Columbia University

Key Words: brain imaging, spatial modeling

In brain imaging studies, region-of-interest (ROI) analysis methods are preferred in some situations to voxel-level analysis because of relative computational efficiency, simpler statistical analysis, and an assurance of comparing "apples with apples" between subjects. However, ROI analysis relies on a seldom validated assumption of homogeneity within a region. We describe general techniques for assessing homogeneity within regions based on dynamic PET data, including both graphical diagnostics and formal testing procedures.

The Use of Multiscale Methods To Characterize Resting fMRI Data

❖Nicole Lazar, University of Georgia, Department of Statistics, Athens, GA 30602, *nlazar@stat.uga.edu*

Current approaches to the analysis of fMRI data assume simple models for the measurement error or for the structure of resting data. These include temporal independence and models for (short or long term) temporal dependence. In the absence of a long time series of data, which will permit the systematic examination of the various models, they remain assumptional in nature. However, statistically valid statements regarding performance of tests require a better characterization of the dependence structure at the voxel level. We investigate the temporal dependence in a long resting fMRI data set; such a time horizon is sufficient to provide insight into possible long-range structure, something that is not feasible with the typical length of scan. Using SiZer and wavelet spectra, we show that voxels in different parts of the brain should be modeled with different dependence structure.

495 Repeated Measures and Recurrent Events in Medical Device Studies ● ✪

Biopharmaceutical Section, WNAR

Thursday, August 2, 8:30 am–10:20 am

Group-Sequential Analysis of Correlated Binary Data on Repeated Venous Accessions

❖John Evans, Boston Scientific Corporation, 100 Boston Scientific Way, Marlboro, MA 01752-1242, *evansj1@bsci.com*; Brian P. Johnson, Boston Scientific Corporation; So Jung Imm, Boston Scientific Corporation

Key Words: Group-sequential design, Repeated measures, Correlated binary data

Evans et al. (2006) gave sample sizes and analyses for correlated binary data across time in a venous access study to compare implantable ports. However, the study also required an interim analysis. Group-sequential designs and analyses for uncorrelated binary data are not appropriate for correlated binary data. Incorrect conclusions may be drawn when comparing proportions of complications. We will present group-sequential methods for correlated binary repeated measures in a venous access study. As for any group-sequential design, sample size is slightly higher than for doing just a final analysis, to allow for the interim analysis to be done, whilst preserving the overall significance level. We will compare results from interim and final analyses under the usual assumption of independence of accessions, with those obtained when a positive correlation is assumed among accessions.

Longitudinal Data Analysis in Medical Device Clinical Studies

❖Chang S. Lao, Food and Drug Administration, 1350 Piccard Drive, Rockville, MD 20850, *Chang.Lao@FDA.HHS.GOV*; Harry F. Bushar, Food and Drug Administration

Key Words: Piecewise Linear spline model, Summary statistics, Repeated measurements, Analysis of covariance

Several statistical approaches are discussed for analysis of longitudinal clinical data from a two independent-group, randomized, prospective, multicenter study of patients with benign prostate hyperplasia (BPH) who are treated with either a new transurethral microwave thermotherapy device (TUMT) or sham control during a 6-month follow-up study. The primary clinical endpoint, the American Urological Association (AUA) symptom score is illustrated in this paper. The primary effectiveness null hypothesis is no treatment effect at any follow-up time against the alternative hypothesis of a favorable treatment effect at one or more follow-up times. The change of baseline and longitudinal slopes were analyzed by a piecewise linear spline regression model with knot at month 1. Statistically significant lower mean AUA scores were found in the TUMT group than the sham group at months 3 and 6.

Comparison of Methods for Analyzing Recurrence Cardiac Arrhythmia Data

❖Shelby Li, Medtronic Inc., 1015 Gramsie Road, Shoreview, MN 55126, *shelby.li@medtronic.com*

Key Words: recurrent events, Cardiac device, Clinical Trials

In cardiac device clinical trials, cardiac arrhythmia episodes are often the center interest of many researchers; therefore they are also chosen as one study end point. Counting process models and analysis of recurrent events models could both be adopted to assess the risk factors for an event. In this exercise, results from fitting a log-linear model, a conditional model, a marginal model and a time-to-first event model are compared.

Assessing Correlation Between Repeated Measurements of Secondary Endpoints in a Medical Device Trial

❖ Lan Pan, Boston Scientific Corporation, One SCIMED Place A242, Weaver Lake Rd, Maple Grove, MN 55311, *lan.pan@bsci.com*; Bryan Randall, Boston Scientific Corporation; Alex Shih, Boston Scientific Corporation

Key Words: Longitudinal Methods, General Linear Model, Structured Covariance, Missing Values

Change in hypertension control is one of the important secondary endpoints in a Boston Scientific Renal Stent System trial. Blood pressures were measured several times during the early and long-term follow-up. In previous analyses of the trial data, only the cross-sectional measurements of blood pressure were used to investigate the change in hypertension control between two baseline resistive index subgroups. In this presentation, we conduct the longitudinal data analysis to assess the correlation between repeated measurements and provide more efficient estimators. We conduct multiple model comparisons to select the most appropriate covariance structure. We also fit the model on balanced and unbalanced data to examine the impact of missing values. The results do not show baseline resistive index as a predictor of blood pressure change in this patient population.

Analysis of Recurrent Events Data in Medical Device Clinical Trials

❖ Sarah J.H. Kogut, Beckman Coulter, Inc., 1000 Hazeltine Drive, Chaska, MN 55318, *sjkogut@beckman.com*

Key Words: clinical trial, medical device, recurrent events, mean cumulative function (MCF), longitudinal data

The safety hypothesis for a medical device clinical trial may explore that subjects treated with the test device have fewer major adverse events during the study period as compared with subjects treated with the control. Adverse events of interest may have repeated occurrences. Time-related accumulation of major adverse events within each treatment group can be estimated using the Nelson nonparametric method. This method is based on graphical comparisons and describes the mean number of major adverse events experienced by subjects over time. The mean cumulative function (MCF) difference between test and control groups can be compared using point wise confidence limits at time of events. Examples of MCF plots and conclusions to be observed will be discussed in the context of a cardiovascular medical device clinical trial.

496 Nonparametric Approaches to Regression and Spatial Modeling ✪

IMS, Section on Nonparametric Statistics, Section on Statistics and the Environment

Thursday, August 2, 8:30 am–10:20 am

Estimating the Innovation Distribution in Nonparametric Autoregression

❖ Anton Schick, Binghamton University, 4032 Marietta Dr, Vestal, NY 13850, *anton@math.binghamton.edu*; Ursula Mueller, Texas A&M University; Wolfgang Wefelmeyer, Univeristaet zu Koeln

Key Words: Residual-based empirical distribution function, local linear smoother, Bahadur representation

We prove a Bahadur representation for a residual-based estimator of the innovation distribution function in a nonparametric autoregressive model. The residuals are based on a local linear smoother for the autoregression function. Our result implies a functional central limit theorem for the residual-based estimator.

2-D Wavelet-Based Spectra with Applications in Analysis of Geophysical Images

❖ Brani Vidakovic, Georgia Institute of Technology, 313 Ferst Drive, Atlanta, GA 30332, *brani@gatech.edu*; Orietta Nicolis, University of Bergamo; Claudio Garutti, University of Padova

Key Words: Wavelets, Self-similarity, 2D wavelet spectra, 2D fractional Brownian Motion

We propose a wavelet-based spectral method for estimating the Hurst parameter in isotropic and anisotropic nonstationary fractional Gaussian fields. The method can be applied to self-similar images and, in general, to d-dimensional data that scale. In the application part, we consider denoising of 2-D fractional Brownian fields and the classification of the clouds/temperature satellite images. In the first application, we use Bayesian inference in the wavelet-spectral domain to separate the signal (the Brownian field) and noise. For the classification of geophysical images we first estimate directional Hurst exponents and use them as an input to standard machine learning algorithms. This work is joint with Orietta Nicolis, University of Bergamo, and Claudio Garutti, University of Padova, Italy.

A Data-Driven Block Thresholding Approach to Wavelet Estimation

❖ Harrison Zhou, Yale University, 24 Hillhouse Ave, New Haven, CT 06511, *huibin.zhou@yale.edu*; Tony Cai, University of Pennsylvania

Key Words: Adaptivity, Block thresholding, James-Stein estimator, Stein's unbiased risk estimate, Wavelets, Nonparametric regression

A data-driven block thresholding procedure for wavelet regression is proposed and its theoretical and numerical properties are investigated. The procedure empirically chooses the block size and threshold level at each resolution level by minimizing Stein's unbiased risk estimate. The estimator is sharp adaptive over a class of Besov bodies and achieves simultaneously within a small constant factor of the minimax risk over a wide collection of Besov Bodies including both the dense and sparse cases. The procedure is easy to implement. Numerical results show that it has superior finite sample performance in comparison to the other leading wavelet thresholding estimators.

Bayesian Nonparametric Modeling for Spatial Data Using Dirichlet Processes

❖ Michele Guindani, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd Unit 447, Houston, TX 77030, *mguindani@gmail.com*

Key Words: Gaussian process, Bayesian Nonparametrics, Nonstationary process, spatial random effects

We provide an overview of recent Bayesian Nonparametric modeling for the analysis of spatial data. The nonparametric specifications that underlie this work are generalizations of Dirichlet process mixture models. We attempt to interrelate these various choices either as generalizations or suitable limits. We offer comparison among these specifications by means of simulation and real data examples.

On Block Thresholding in Wavelet Regression with Long Memory Correlated Noise

❖ Linyuan Li, University of New Hampshire, Department of Mathematics and Statistics, Durham, NH 03824, *linyuan@math.unh.edu*

Key Words: Block thresholding, long range dependence data, minimax estimation, rates of convergence

We consider block thresholded wavelet-based estimators of mean regression functions with long memory Gaussian errors and investigate their asymptotic convergence rates. We show that these estimators, based on block thresholding of empirical wavelet coefficients, achieve optimal minimax convergence rates over a large range of Besov function classes. These results are analogous to those in the standard regression problem with i.i.d. Gaussian errors. Therefore, in the presence of long memory Gaussian errors, wavelet estimators still provide extensive adaptivity over a wide range of Besov classes.

497 Bayesian Modeling of High-Dimensional Genomic Data ● ✪

Section on Bayesian Statistical Science, ENAR, Biometrics Section

Thursday, August 2, 8:30 am–10:20 am

Using Genomic Scans of Fst To Identify the Footprints of Selection

❖ Kent Holsinger, University of Connecticut, Dept of Ecology and Evolutionary Biology, Storrs, CT 06269-3043, *kent@darwin.eeb.uconn.edu*

Key Words: Fst, hierarchical models, genetic diversity, natural selection, genomics

Wright's Fst is widely used to describe the partitioning of genetic variation within and among populations. Estimates of Fst at all genetic loci sampled from a set of populations will be statistically indistinguishable, unless mutation rates differ substantially across loci or unless loci are subject to different patterns of natural selection. I will present a hierarchical Bayesian model that provides estimates of locus- and population-specific effects on Fst, and I will illustrate how it can be applied to the analysis of high-resolution genomic data. Analysis of these data introduces a novel approach to outlier detection. Analysis of the high-resolution data will also introduce a novel application of spatial modeling in genetic data.

Bayesian Semiparametric Testing and Estimation for High-Dimensional Genetic Data

❖ Richard MacLehose, National Institute of Environmental Health Sciences, PO 12233, MD A303, Research Triangle Park, NC 27709, *maclehoser@niehs.nih.gov*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Dirichlet Process, SNP, Bayes

Genotyping technology has made analysis of large dimensional data commonplace, with the number of predictors typically exceeding the number of observations. Statistical methods are needed to test associations between a vast number of SNPs and a phenotypic outcome, while accounting for high levels of between SNP correlation. False discovery rates are often used in this setting to control the expected proportion of false positives, however they typically ignore the problem of false negatives. We adopt a decision theoretic Bayesian approach which penalizes false positives and false negatives. To reduce dimensionality, we focus on the clustering properties of the Dirichlet Process (DP). We adopt a "retrospective" approach to estimating DP models that is more robust to high dimensions. We discuss various procedures for hypothesis testing and compare our method to FDR-based methods.

Bayesian Pathways Studies Using Microarray Data

❖ Lynn Kuo, University of Connecticut, 215 Glenbrook Rd, Storrs, CT 06269-4120, *lynn@stat.uconn.edu*; Wangang Xie, University of Connecticut; Dong-Guk Shin, University of Connecticut; Fang Yu, University of Connecticut; Yifang Zhao, University of Connecticut; Ming-Hui Chen, University of Connecticut

Key Words: model selection, Bayes factor, MCMC, microarray data, Bayesian network, molecular pathways

Identifying molecular pathways that are most activated in a defined stage of cell differentiation or when cells are exposed to environmental stimuli provides more insights on the functional information about genes. We propose novel methods to evaluate a set of possible pathways obtained from KEGG or BioCarta data bases on their activations from the microarray and proteomic data. These high-throughput data are further supplemented by prior information that is constructed from literature search on the gene-to-gene promotion or inhibition knowledge. The Bayes factor approach is used to evaluate the evidence for each activated pathway. Essentially, we develop Markov chain Monte Carlo methods and Bayesian model selection methods to identify a set of pathways that are most activated by observing the high-throughput data.

Bayesian Graph Cutting

❖ Bani Mallick, Texas A&M University, Department of Statistics, 3143, College Station, TX 77843-3143, *bmallick@stat.tamu.edu*

Key Words: Microarray, gene networks, graph cutting, Dirichlet process, kernel k-mean

Graph cutting techniques prevail in the non-statistical community for identifying meaningful subgraphs or significantly interacting subset of vertices in a large graph. We explore the nonparametric Bayesian partitioning of graphical structures representing the dependency between different responses in a dataset. We propose a nonparametric Bayesian alternative to graph cutting by approaching it as a clustering problem. For clustering, we use nonparametric Dirichlet process priors . It is shown that loss functions similar to the kernel k-means naturally arises in this model and the minimization of associated posterior risk comprises an effective graph cutting strategy. The method has been applied to two microarray datasets namely, the melanoma dataset (Bittner et al., 2000) and the sarcoma dataset (Nykter et al., 2006). Some preliminary results are reported.

Log-Linear Models for Gene and Network Interactions

❖ Valen Johnson, The University of Texas M.D. Anderson Cancer Center, Biostatistics and Applied Mathematics, Houston, TX , *vejohnson@mdanderson.org*; Jianhua Hu, The University of Texas M.D. Anderson Cancer Center

Key Words: Bayesian model selection, gene interactions, network analysis, log-linear model

We describe a Bayesian approach to the analysis of gene and network interaction data based on log-linear models. The goal of our approach is inference regarding causal connections among genes and functional pathways, and is based on a Bayesian model selection algorithm applied to data that has been reduced to ordered categories using ranks of observations within both subjects and genes or network components.

498 Bayesian Methods in Industrial Applications ● ✪

Section on Bayesian Statistical Science

Thursday, August 2, 8:30 am–10:20 am

A Bayesian Approach to Assessing the Risk of QT Prolongation

❖Suraj Anand, North Carolina State University, , *suraj.anand@gmail.com*; Sujit Ghosh, North Carolina State University

Key Words: QT prolongation, QTc, Monte Carlo simulation, Bayesian

The standard approach to investigating a drug for its potential for QT prolongation is to construct a 90% two-sided (or a 95% one-sided) confidence interval, for the difference in mean QTc (heart-rate corrected version of QT) between drug and placebo at each time point, and to conclude noninferiority if the upper limits for all these CIs are less than a prespecified constant. One alternative approach is to base the noninferiority inference on largest difference in population mean QTc between drug and placebo. In this paper, we propose a Bayesian solution to this problem using a Monte Carlo simulation method. The proposed method has several advantages over some of the other existing methods and is easy to implement in practice. We use simulated data to assess the appropriateness of this approach and apply the method to a real dataset on QTc to make statistical inference.

Bayesian Nonparametric Random Effects Selection Using the Centered Dirichlet Process

❖Mingan Yang, National Institute of Environmental Health Sciences, NIH, 3950 Jackson St Apt H24, Raleigh, NC 27607, *yangm2@niehs.nih.gov*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: nonparametric Bayes, Dirichlet process, mixed model, model uncertainty, Stochastic Search, Variable selection

Random effects models are widely used to account for dependence in repeated observations, collected on a subject over time or for different subjects in a center. In conducting inferences and building predictive models, there is typically uncertainty in the subsets of predictors to be included in the fixed and random effect components, as well as in the distribution of the random effects. A challenge in nonparametric modeling of random effects distributions is the need for zero mean constraints to avoid bias. This article proposes an approach for subset selection in nonparametric models. By using a centered Dirichlet process mixture of Gaussians, we choose a prior for the random effects density which has support on the space of absolutely continuous densities with zero mean. An efficient parameter-expanded stochastic search Gibbs sampler is developed, which allows inferences.

Bayesian Estimation of Scram Rate Trends in Nuclear Power Plants

❖Kaushal Mishra, North Carolina State University, 1707 Crest Road Apt 4, Raleigh, NC 27606, *kkmishra@gmail.com*; Sujit Ghosh, North Carolina State University

Key Words: Bayesian Inference, MCMC, Zero-inflated Model

Nuclear reactors are equipped with reactor scram systems to ensure rapid shutdown of the system in the event of leaks, failure of power conversion systems, or other operational abnormalities. The U.S. Nuclear Regulatory Commission (NRC) collects data of scram rate for various nuclear power plants to obtain their trend of proper functioning over time and to regulate them. The annual scram data obtained from 66 commercial nuclear power plants indicate an increase in no scram event from 1.5% in 1986 to 33% in 1993. To analyze such a correlated count data with excess zeros, a zero-inflated model that accounts for both temporal and plant to plant variation is being proposed. A wide class of possibly non-nested models was fitted using MCMC methods and compared using a predictive criterion.

The Use of Bayesian Tolerance Intervals for Batch-Release Testing

❖Richard Lewis, GlaxoSmithKline, PO Box 13398, Five Moore Drive, Research Triangle Park, NC 27709-3398, *richard.a.lewis@gsk.com*; Steven J. Novick, GlaxoSmithKline

Key Words: Tolerance Intervals, Random-Effects Models, Batch-Release Testing, Bayesian Tolerance Intervals

When developing quality-control criteria such as batch-release limits, one would like to have a high level of confidence that acceptable batches will meet the criteria. Tolerance limits seem well-suited for this purpose and their use has previously been proposed in a number of settings, ranging from simple random samples from a Gaussian population to random-effects models. We consider using Bayesian tolerance intervals to determine batch-release limits and simulate their performance in some selected settings, including random-effects models with unbalanced data.

A Joint Modeling Approach for Analyzing Nonignorable Missing Data

❖Sujit Ghosh, North Carolina State University, 2501 Founders Drive, Department of Statistics, Raleigh, NC 27695-8203, *ghosh@stat.ncsu.edu*; Liansheng Zhu, Pharmaceutical Product Development, Inc

Key Words: longitudinal data, joint modeling, mcmc, nonignorable missing, clinical trial

In longitudinal studies, data are often missing and when the missing mechanism is informative and hence not ignorable, it is generally difficult to analyze such non-ignorable missing (NIM) data since the distributional assumptions about missing data are not easily verifiable. Within the class of pattern-mixture models we develop a joint-modeling (JM) approach, in which patterns considered as random effects are marginalized within a generalized linear mixed model framework. The JM approach is shown to be able to capture the dependence of missing indicators on missing outcomes in some degree as is the case with NIM data. Some of the main advantages of the proposed approach include (i) the capability to handle both continuous and discrete responses, (ii) avoidance of the problem of under-identifiability, (iii) availability of marginal estimates, and (iv) computational efficiency.

499 Ranked Set Sampling I ✪

Section on Nonparametric Statistics, Section on Survey Research Methods

Thursday, August 2, 8:30 am–10:20 am

Estimating the Distribution Function Using k-tuple Ranked Set Samples

❖Kaushik Ghosh, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, *ghosh@njit.edu*; Ram C. Tiwari, National Cancer Institute

Key Words: Extreme ranked set sample, Imperfect ranking, Cost of sampling, Empirical process

The basic assumption underlying the concept of ranked set sampling is that actual measurement of units is expensive whereas ranking is cheap. This may not be true in reality in certain cases where ranking may be moderately expensive. In such situations, based on total cost considerations, k-tuple ranked set sampling is known to be a viable alternative, where one selects k units (instead of one) from each ranked set. In this article, we consider estimation of the distribution function based on k-tuple ranked set samples when the cost of selecting and ranking units is not ignorable. We investigate estimation both in the balanced and unbalanced data case. Properties of the estimation procedure in the presence of ranking error are also investigated. Results of simulation studies as well as an application to a real data set are presented to illustrate some of the theoretical findings.

Distribution-Free Statistical Intervals via Ranked-Set Sampling

❖Jesse Frey, Villanova University, Department of Mathematical Sciences, 800 Lancaster Ave, Villanova, PA 19085, *jesse.frey@villanova.edu*

Key Words: confidence intervals, prediction intervals, tolerance intervals

This paper discusses the use of ranked-set sampling to create distribution-free statistical intervals. The intervals developed include confidence intervals, tolerance intervals, and prediction intervals. Methods for computing coverage probabilities are derived, and the performance of the intervals is compared to that of intervals based on simple random sampling. The ranked-set-sampling-based intervals are found to be at least as good in most settings of interest, but the nature of the advantage depends on the type of interval considered.

A Nonparametric Mean Estimator for Judgment Post-Stratified Data

❖Xinlei Wang, Southern Methodist University, 3225 Daniel Avenue, Dallas, TX 75275-0332, *swang@smu.edu*; Johan Lim, Yonsei University; S. Lynne Stokes, Southern Methodist University

Key Words: Imperfect Ranking, Imprecise Ranking, Isotonic Regression, Multiple Rankers, Ranked Set Sampling, Simple Stochastic Ordering

MacEachern, Stasny and Wolfe (2004) introduced a data collection method, called judgment post-stratification (JP-S), based on ideas similar to those in ranked set sampling, and proposed methods for mean estimation from JP-S samples. In this paper we propose an improvement to their methods, which exploits the fact that the distributions of the judgment post-strata are often stochastically ordered, so as to form a mean estimator using isotonized sample means of the post-strata. This new estimator is strongly consistent with similar asymptotic properties to those in MacEachern et al. (2004). It is shown to be more efficient for small sample sizes, which appears to be attractive in applications requiring cost efficiency. Further, we extend our method to JP-S samples with imprecise ranking or multiple rankers. The performance of the proposed estimators is examined on three data examples.

Optimal Ranked Set Sampling Estimation Based on Medians from Multiple Set Sizes

❖Nader Gemayel, The Ohio State University, 404 Cockins Hall, 1958 Neil Ave, Columbus, OH 43210, *nader@stat.ohio-state.edu*; Elizabeth A. Stasny, The Ohio State University; Douglas Wolfe, The Ohio State University

Key Words: Best linear unbiased estimator (BLUE(, Beta distribution, Lagrange multipliers, order statistics, robust estimation, Taylor's theorem

Ranked set sampling (RSS) is a data collection technique that makes use of expert knowledge to rank sample units before measuring them. Even though rankings are not always perfect, RSS is useful in situations when obtaining measurements is costly, difficult, or destructive. Research in this area has tended to focus on the case of balanced RSS, in which all set sizes are equal and exactly one observation of each rank is measured in each cycle. This article represents a departure from balanced RSS when we encounter different set sizes within a single sample. More specifically, we propose a distribution-free estimator for the median of a symmetric distribution using medians of ranked set samples of various set sizes from this distribution. This estimator is seen to be robust over a wide class of symmetric distributions.

Minimum Distance Estimator of Judgment-Ranking Probabilities in a Ranked-Set Sample

❖Omer Ozturk, The Ohio State University, Department of Statistics, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *omer@stat.ohio-state.edu*

Key Words: imperfect ranking, doubly stochastic, best linear unbiased estimator, calibration, doubly stochastic matrix, perfect ranking

In ranked set sampling procedure, ranking error is almost an unavoidable reality. Thus, a reasonable statistical inference should address the validity of the procedure under imperfect ranking. There are several models that describe the ranking mechanism. One of these models is introduced by Bohn and Wolfe(1994) and later extended by Frey(2006). Their model constructs the judgment class distributions as a mixture distributions of the actual order statistics. In this talk, we first develop a minimum distance estimator for the mixing parameters of this mixture distribution. Then we use this estimated model to calibrate the estimator and tests to reduce effects of ranking errors on statistical procedures.

500 Creating Materials for an Online Course

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences, WNAR

Thursday, August 2, 8:30 am–10:20 am

Developing and Using Video Instruction in an Online Course

❖Sue Schou, Idaho State University, PO Box 4043, Pocatello, ID 83205, *schosue@isu.edu*

Key Words: online, video, instruction

Using the ADDIE model of instructional design, this instructor produced an online introductory business statistics course for the College of Business at Idaho State University. The course was designed with both the learner and the COB course requirements in mind. In the development phase, the instructor created digital video instruction for streaming over the Internet as well as video instruction for using Minitab. Breeze presentations were developed for use as an alternative video format. In addition, the COB purchased a site license for video streaming the Against All Odds series. Using these tools as well as other materials, the students enrolled in the online section of introductory business statistics have been as successful as students taught in a traditional face-to-face setting as evaluated using both the final examination and overall success rate for the course.

Creating Materials for an Online Course

❖Suzanne Perumean-Chaney, The University of Alabama at Birmingham, RPHB 327J, 1530 3rd Ave S, Birmingham, AL 35294, *schaney@uab.edu*; Mischelle T. Stone, The University of Alabama at Birmingham

Key Words: Education, Statistics, On-Line, Undergraduates

This presentation describes the experience of teaching a first-time, online undergraduate statistics course in the Department of Justice Sciences. The course is oriented towards a consumer of statistics course. The course format includes PowerPoint slides with voice-over lectures, homework assignments, and a discussion assignment. The course is assessed with a pre-post knowledge exam, a questionnaire identifying computer use habits and basic demographics, and standard mid-term and final course evaluations. The results in terms of course material preparation, student success, and challenges faced will be discussed.

Creating Materials for Online Students

❖ Patti Collings, Brigham Young University, 231 TMCB, Provo, UT 84602, *collingsp@stat.byu.edu*

Key Words: Introductory statistics, online, education

We have almost one thousand on campus students taking introductory statistics online with another eight hundred or so taking introductory statistics as an independent study course. To prepare materials for these students, we first created lessons using Macromedia Flash. This environment allowed us to animate many statistical concepts. Then, we actually wrote scripts, recorded them and added them to the Flash lessons in QuickTime movies. In this talk, I will describe in more detail how we did this and demonstrate these materials.

Creating Materials for an Online/Distance Course

❖ F. Michael Speed, Texas A&M University, 9312 Lake Forest Ct South, College Station, TX 77845, *mspeed@tamu.edu*; Kim Ritchie, Texas A&M University

Key Words: online education, distance education

If we assume that our on-campus classes are exemplary in providing great learning environments, then it is natural that we try to replicate as close as possible the on-campus experience in our on-line classes. We agree with Oblinger and Hawkins (2006) that learning occurs "as a result of motivation, opportunities, an active process, interaction with others, and the ability to transfer learning to a real world situation." This paper will review the technologies and materials which can lead to the desired learning outcomes and do so with a limited about of expenditures. Diana G. Oblinger and Brian L. Hawkins, "The Myth about No Significant Difference," Educause Review - November/December 2006.

Creating Materials for an Online Course

❖ Joan Weinstein, Harvard Extension School, 14 Hobart Road, Newton Centre, MA 02459, *weinst2@fas.harvard.edu*

Key Words: distance education, online courses

Each semester I teach a large introductory statistics course which includes both local and distance students. Each lecture is videotaped and available online within 48 hours for viewing by any enrolled student at any time during the semester. I have tried to design the course so that the experience for distance learners approximates that of students who are able to attend class. In this presentation I will discuss the techniques I am using to create class materials, communicate with distance students, and address difficulties we have encountered. I will also include feedback from both local and distance students.

501 Bayesian Methods in Sample Surveys

Section on Survey Research Methods
Thursday, August 2, 8:30 am–10:20 am

Parametric and Semiparametric Model-Based Estimates of the Finite Population Mean for Two-Stage Cluster Samples with Item Nonresponse

❖ Roderick J. Little, University of Michigan, Department of Biostatistics, School of Public Health, 1420 Washington Heights, M4208 SPHII, Ann Arbor, MI 48109, *rlittle@umich.edu*; Ying Yuan, The University of Texas M.D. Anderson Cancer Center

Key Words: multiple imputation, missing data, survey data, Bayesian methods

This talk concerns item nonresponse adjustment for two-stage cluster samples. Specifically, we focus on Bayesian multiple imputation for two types of nonignorable nonresponse: nonresponse depending on covariates and underlying cluster characteristics, and depending on covariates and the missing outcome. In these circumstances, standard weighting and imputation adjustments are liable to be biased. To obtain consistent estimates, we extend the standard random-effects model by modeling these two types of missing data mechanism. We also propose semiparametric approaches based on fitting a spline on the propensity score, to weaken assumptions about the relationship between the outcome and covariates. These new methods are compared with existing approaches by simulation. The Behavioral Risk Factor Surveillance System (BRFSS) data are used to illustrate these approaches.

Bayesian Penalized Spline Model-Based Estimation of the Finite Population Distribution Function for Unequal Probability Samples

❖ Qixuan Chen, University of Michigan, 1984 Traver Rd, Apt 204, Department of Biostatistics, Ann Arbor, MI 48105, *qixuan@umich.edu*; Michael R. Elliott, University of Michigan; Roderick J. Little, University of Michigan

Key Words: penalized spline regression, Gibbs sampling, probit model, distribution function, unequal probability samples

This paper develops Bayesian estimation of finite population distribution function for unequal-probability-of-selection samples. The methods allow the probabilities of inclusion to be directly incorporated into the estimation of distribution functions, using an ordinal probit regression or a sequence of binary probit regressions on the penalized spline of the inclusion probabilities. The posterior distribution of the distribution function is then obtained using Gibbs sampling. The proposed methods are compared with the Horvitz-Thompson estimator by simulation, and illustrated using the University of Michigan Dioxin Exposure Study (UMDES) data.

Simultaneous Credible Intervals for Small-Area Estimation Problems

❖ Nadarjasund Ganesh, University of Maryland, 4311 Mathematics Building, Department of Mathematics, College Park, MD 20742, *ganesh@math.umd.edu*

Key Words: Bayesian multiple comparison, EBLUP, hierarchical Bayes, matched priors, mixed linear models, Monte Carlo

In this paper, we fill in an important research gap in the small area literature, namely the problem of multiple comparison. We illustrate how the Bayesian approach can be applied to develop different multiple comparison procedures. Unlike the classical methods, the Bayesian method can incorporate all sources of uncertainties, irrespective of the sample size, and can make inferences conditional on the data. However, the selection of priors for the hyperparameters is one of the important problems in Bayesian analysis. We revisit this problem in the context of multiple comparison, and develop a new class of objective priors that approximate certain well-known frequentist's solutions. This class includes the well-known superharmonic prior due to Stein. The utility of our method is illustrated through data analysis.

Bayesian Approaches to Sequential Selection of Survey Design Protocols

❖James Wagner, University of Michigan, G394 Perry, PO Box 1248, Ann Arbor, MI 48106-1248, *jameswag@isr.umich.edu*; Trivellore E. Raghunathan, University of Michigan

Key Words: Adaptive Design, Survey Design Protocols

Surveys sequentially choose design protocols for contacting and interviewing cases. Most surveys follow very general protocols that minimally differentiate cases from each other. We develop a series of models to help identify most efficient strategies for cases conditional on their fixed characteristics and history of previous attempts. These models use priors specified from other surveys and data from the sampling frame and call records to derive posterior probabilities of contact and interview. These posterior probabilities are then used to determine the best protocol to be used for maximizing the chances of contact or interview. Results from random digit dial and face-to-face surveys analyzed from this perspective are presented.

502 Methods for Inference about Population Structure and Inheritance from Genetic Data ● ✪

WNAR, Section on Statistics in Epidemiology, Biometrics Section
Thursday, August 2, 8:30 am–10:20 am

A Likelihood Method of Estimating Ancestry from Pedigrees

❖Janet Sinsheimer, University of California, Los Angeles, Human Genetics 5357C Gonda, David Geffen School of Medicine, Los Angeles, CA 90095, *janet@mednet.ucla.edu*; Christopher Plaisier, University of California, Los Angeles; Adriana Huertas-Vazquez, University of California, Los Angeles; Carlos Aguilar-Salinas, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán Mexico City; Teresa Tusie-Luna, National Autonomous University of Mexico; Paivi Pajukanta, University of California, Los Angeles; Kenneth Lange, University of California, Los Angeles

Key Words: ethnic admixture, statistical genetics, likelihood, population structure, estimation, genomics

We have developed a new likelihood based method of estimating ancestry admixture that uses unrelated individuals or pedigrees. For each founder of a pedigree, the method estimates the likelihood of ancestry in K ancestral populations based on the observed genotypes at ancestry informative markers throughout his/her pedigree. These estimates are then propagated down the pedigree to the nonfounders. By using a likelihood based approach, standard errors of the founders' proportions can be determined and propagated to nonfounders. An individual's ethnic ancestry can be more precisely estimated by using the genotype information available in other family members than compared to just using the individual's genotypes alone.

Cluster Analysis and PCA for Detection of Genetic Variants and Population Structure

❖Diana Luca, Carnegie Mellon University, 4910 Centre Ave, Ap F2, Pittsburgh, PA 15213, *dluca@stat.cmu.edu*; Kathryn Roeder, Carnegie Mellon University; Bernie Devlin, University of Pittsburgh Medical Center

Key Words: case-control, genetic markers, cluster analysis, PCA, population heterogeneity

Case-control studies for association are widely used for finding genetic variants causally associated with phenotypes. Unfortunately, any type of population structure can induce false positives. If cases and controls have different genetic backgrounds, there is a chance that some gene frequency differences are due to ancestral population of origin rather than to the marker loci's direct cause of disease. This study focuses on the close relationship between cluster-based methods and PCA techniques. Simulation results show that the two approaches can complement each other in order to improve the real structure detection.

Evaluating Individual Admixture Estimation Through Plasmodes

❖Laura K. Vaughan, The University of Alabama at Birmingham, RPHB 327, 1530 3rd Ave S, Birmingham, AL 35294-0022, *lkvaughan@uab.edu*; Jasmin Divers, The University of Alabama at Birmingham; Miguel Padilla, The University of Alabama at Birmingham; David T. Redden, The University of Alabama at Birmingham; Hemant Tiwari, The University of Alabama at Birmingham; Daniel Pomp, The University of North Carolina at Chapel Hill; David B. Allison, The University of Alabama at Birmingham

Key Words: Admixture, Ancestry, Plasmode, Structured Association Testing, Regional Admixture Mapping

Structured association testing, where individual admixture estimates are used as control variables, has been advanced as a means of controlling for confounding induced by population stratification in genetic association studies. Many admixture estimation methods have been proposed and evaluated with simulated data. An attractive method to supplement the evaluation is the use of plasmodes. Plasmode datasets are simply those generated by natural processes, and not through computer simulation, but have some aspect of the truth known. Here we use plasmodes consisting of mouse crosses. Each cross shares a common founding inbred line so that the ancestry proportion for each mouse is known. This allows us to compare the effectiveness of different algorithms (Structure, AdmixMap, and FRAPPE) at estimating individual admixture proportions.

Population Structure and Covariate Analysis Based on Pairwise Microsatellite Allele Matching Frequencies

❖Geof Givens, Colorado State University, Dept of Statistics, 1877 Campus Delivery, Fort Collins, CO 80523, *geof@lamar.colostate.edu*; Isin Ozaksoy, Colorado State University

Key Words: genetics, generalized additive model, biology, permutation test

We describe a general model for pairwise microsatellite allele matching probabilities. The model can be used for analysis of population substructure, and is particularly focused on relating genetic correlation to measurable covariates. The approach is intended for cases when the existence of subpopulations is uncertain and a priori assignment of samples to hypothesized subpopulations is difficult. The model estimates genetic structure associated with a Wahlund effect and also with spatial, temporal, or other variables that may confound the detection of population structure.

Hypothesis tests for population substructure and for covariate effects can be carried out using permutation methods. Simulated and real examples illustrate the effectiveness and reliability of the approach. The method is compared to other familiar approaches.

Uncertainty in Inheritance: P-Values and Confidence Levels

❖ Elizabeth Thompson, University of Washington, Dept Statistics, Box 354322, Seattle, WA 98195-4322, *thompson@stat.washington.edu*

Key Words: genetic linkage, pedigree data, inheritance patterns, genome-wide mapping, multiple tests, latent variables

In any latent variable problem, in which the ideal test statistic is a function of latent variables S, one may consider the p-value p(S). The latent p-value is then defined as the probability distribution of p(S) given the observed data Y. A randomized test and confidence sets for the parameter of interest follow. Here we apply this idea to testing for genetic linkage for a quantitative trait, and to estimating the location of DNA affecting a trait, given genetic marker data on a pedigree structure. The approach provides not only an exact p-value as a measure of significance, including a natural correction for multiple testing over the genome, but also a measure of the uncertainty associated with that p-value given that S is not observable. Confidence sets are a natural extension that reveal the asymmetry between absence and presence of linkage.

503 Survival Data Analysis II ● ✪

Biometrics Section

Thursday, August 2, 8:30 am–10:20 am

Modeling Bivariate Survival Times by Frank Copula

❖ Rui Qin, Mayo Clinic, Division of Biostatistics, 200 First Street SW, Rochester, MN 55905, *qin.rui@mayo.edu*; Michael P. Jones, The University of Iowa

Key Words: bivariate survival times, copula, pseudolikelihood, negative association

Copula models with relative risk margins are proposed for bivariate survival time data. The comprehensive family of Archimedean copulas provides flexibility in modeling different correlations. Estimation occurs in two stages. First, Cox regression is used to estimate the marginal survival functions. Second, a pseudo-likelihood for the association parameter is constructed by plugging in the marginal estimators and is then maximized over the association parameter. Empirical process theory is applied to establish consistency and asymptotic normality of the two-stage estimator. Simulation is used to study the behavior of the estimator and the asymptotic conclusions in small to medium size samples. The Frank copula exhibits some advantages over the popular Layton copula. An example with a real dataset is given.

Nonparametric Inference for Clustered Competing Risks Data

❖ Bingshu Chen, National Institutes of Health, 6120 Executive Blvd, Rockville, MD 20851, *cheneric@mail.nih.gov*; Philip S. Rosenberg, National Institutes of Health; Joan L. Kramer, National Institutes of Health; Mark H. Greene, National Institutes of Health

Key Words: Competing risks, Correlated survival data, Counting processes, Nonparametric estimate, Two-sample test

Clustering arises naturally in clinical genetic studies and other settings. We develop a nonparametric estimator of cumulative incidence, and obtain robust pointwise standard errors, that account for within-cluster correlation. We modify the two-sample Gray and Pepe-Mori tests for correlated competing risks data, and propose a simple two-sample test of the difference in cumulative incidence at a landmark time. In simulation studies, our estimators are asymptotically unbiased, and the modified test statistics control the type I error. We apply our methods to a family-based prospective cohort study of hereditary breast/ovarian cancer families. For women with BRCA1 mutations, we estimate the cumulative incidence of breast cancer in the presence of competing mortality from ovarian cancer, accounting for significant within-family correlation.

Pseudospline-Based Flexible Modeling of Survival Data

❖ Zekarias Berhane, Drexel University, 1505 Race St Bellet Bldg 1139, School of Public Health Drexel U, Philadelphia, PA 19119, *zb26@drexel.edu*; Lisa Weissfeld, University of Pittsburgh; Chung-Chou H. Chang, University of Pittsburgh

Key Words: Survival, Pseudospline, Spline, Hazards, Proportional

In 1996, Hastie used pseudo-eigenvalue decomposition to approximate the smoothing matrix in an additive model setting based on orthonormal basis functions. In developing a new model, we extended the pseudospline approach to the multiplicative setting proposed by Hastie and Tibshirani (1990) and expanded on Cox's 1972 proportional hazards model. The use of pseudosplines in models of survival data offers several advantages, including flexibility in the specification of the form of the hazard function, a reduction in the computational burden of the estimation imposed by Gray's 1992 model of survival analysis, and simplification of model inferences. We will describe our pseudospline-based model and examine its results when we applied it to the analysis of survival data.

Accommodating Deaths in Longitudinal Self-Rated Health for Hematopoietic Stem Cell Transplant Recipients

❖ Brenda F. Kurland, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., D5-360, P.O. Box 19024, Seattle, WA 98109-1024, *bkurland@fhcrc.org*; Paul J. Martin, Fred Hutchinson Cancer Research Center; Stephanie J. Lee, Fred Hutchinson Cancer Research Center

Key Words: competing risks, longitudinal data, self-rated health

A patient's health status in the years following stem cell transplantation may be excellent, or may be marred by toxicity, late effects, or death. When modeling health status over time, inclusion or exclusion of patients who die may lead to overly pessimistic or optimistic projections. Three approaches will be compared for measuring health status in 320 patients up to 4 times within 2 years after transplant: 1) probability of being healthy and alive, a joint model of health and survival in the entire cohort; 2) health status of survivors at each timepoint, a partly conditional model describing health status in the dynamic cohort of survivors; and 3) description of observed health trajectories for individuals, a pattern-mixture model. Most candidate models are fitted using standard software. Predicted values and group differences may differ greatly based on modeling choices.

Variable Importance in Highly Predictive Models

❖ Kyle Rudser, University of Washington, Biostatistics, Box 357232 1705 NE Pacific, Seattle, WA 98195-7232, *rudserk@u.washington.edu*; Michael LeBlanc, Fred Hutchinson Cancer Research Center; Scott Emerson, University of Washington

Key Words: Variable importance, Predictive models, Nonparametric, Survival Analysis

The hazard ratio is commonly used for comparing survival distributions. While easily estimated in the presence of censored data, it does not allow for the clinical relevance of differences in survival across groups to be easily judged. We consider an approach to nonparametric inference for clinically meaningful functionals of a survivor distribution (e.g., restricted mean, quantiles). In this approach we use different models to borrow information across sparse data than to form contrasts. Linear contrasts are evaluated and compared on mean squared error between approaches using nonparametric recursive partitioning, Cox's proportional hazards, and Buckley-James' linear regression with censored data. The nonparametric approach was superior to the semiparametric approaches when semiparametric model assumptions were violated, and had a slight loss of efficiency when they do hold.

Low-Structure Discrete Hazard Estimation by Cross-Validated Modified Likelihood

❖Steven Mongin, The University of Minnesota, 200 Oak Street SE Suite 350, Minneapolis, MN 55455-2008, *sjmongin@umn.edu*; Timothy R. Church, The University of Minnesota Twin Cities

Key Words: hazard estimation, kernel smoothing, bootstrap, cross-validated modified likelihood

Given a set of survival data, the problem of hazard estimation and inference with little prior knowledge of structure is common. Standard estimators in the non- or semi-parametric category are often employed as a first step to identifying a final model which fits the data–perhaps too well, resulting in overfitting. On the other hand, the flexibility of fully nonparametric estimators may be more than necessary, given a premise of underlying continuity in the hazard under investigation. As a remedy, a method of hazard estimation is described which is related to kernel smoothing and features automatic bandwidth selection. Cross-validated modified likelihood (CVML) is invoked, with adaptations for the problem of discrete hazard estimation in potentially large data sets, using the stratified bootstrap. Applications include CRC incidence in the Minnesota Colon Cancer Control Study.

Development of a Stochastic Model of Salmon Migration Using Random Effects Models of Recapture Data from Tagged Individuals

❖Steven G. Smith, NOAA Fisheries, NWFSC US Dept. of Commerce, 2725 Montlake Blvd E, Seattle, WA 98112, *steven.g.smith@noaa.gov*; James R. Faulkner, National Marine Fisheries Service; Richard W. Zabel, NOAA Fisheries

Key Words: random effects models, capture-mark-recapture, stochastic modeling, fisheries management, variance components

COMPASS (COMPrehensive PASSage) is a new model of the migration of juvenile salmon migrating through the hydroelectrically developed Snake and Columbia rivers of the northwestern United States. Model relationships between environmental conditions and dam operations and the travel time and survival of migrating fish have been developed using data from electronically tagged fish. To use the model as an effective management tool it must generate both predictions about fish response to management actions and measures of uncertainty about those predictions. The underlying tag data vary both spatially as fish move downstream and temporally, both within and across years. The variation has components of both sampling error and process error. We use random effects models to disentangle the sources of variation, so that prospective model predictions appropriately reflect uncertainty.

504 ROC Estimation and Diagnostic Tests ● ✪

Biometrics Section

Thursday, August 2, 8:30 am–10:20 am

Statistical Evaluation of Diagnostic Accuracy in Multiple Diagnostic Markers

❖Mi-Chia Ma, National Cheng Kung University, No. 1 Ta-Hsueh Rd., Department of Statistics, Tainan, 70101 Taiwan, *mcma@stat.ncku.edu.tw*; Kun-Cheng Lin, National Cheng Kung University

Key Words: diagnostic accuracy, the area under multivariate ROC curve, nonparametric estimation, the random effect model

This paper studies the problem of estimating the diagnostic accuracy when several diagnostic markers are used to detect a disease. The area under multivariate ROC curve or the volume under multivariate ROC surface is used to measure the diagnostic accuracy. A nonparametric estimator of the diagnostic accuracy and its variance are proposed. A model-based estimate of the diagnostic accuracy and its variance are also derived based on the random effect model. The model-based estimate is compared with the nonparametric estimator by the mean square error. The approximated confidence intervals are derived and evaluated by the nominal coverage probability based on a simulation study. The simulation results show that the mean square error of the model-based estimator is smaller than that of the nonparametric estimator. Finally, the proposed method is applied to a real data for example.

Additional Diagnostic Utility of a Variable After the Adjustment of Other Variables

❖Caixia Li, University of California, San Francisco, Dept of Radiology Box0946, 185 Berry Street Suite 350, San Francisco, CA 94143-0946, *caixia.li@radiology.ucsf.edu*; Ying Lu, University of California, San Francisco

Key Words: ROC curves, Linear Discriminant Function, Linear Transformation, Type I error, Multivariate F-distribution

Diagnostic information from multiple sources is common in medicine. The additional diagnostic value of a new variable after the adjustment of other variables can be evaluated by the incremental area under (AUC) the receiver operating characteristic (ROC) curves. However, such a statistical testing procedure fails to maintain proper type I error rates for practically used sample sizes. In this paper, we investigated a new test procedure based on a standardized mean difference between the location parameters of two ROC curves after a linear transformation of all diagnostic variables. We used our method to test hypothesis of superiority or non-inferiority. We proved, through simulation studies, that such a test can maintain better type I error rates as well as unbiased estimate of the standardized effect sizes. We used an example of hip fracture to illustrate our method.

Adjusting for Covariate Effects on Classification Accuracy Using the Covariate-Adjusted ROC Curve

❖Holly Janes, Johns Hopkins University, 615 N Wolfe St, Department of Biostatistics, Baltimore, MD 21205, *hjanes@jhsph.edu*; Margaret Pepe, Fred Hutchinson Cancer Research Center

Key Words: classification accuracy, sensitivity, specificity, ROC curve, covariate adjustment

In evaluating markers for disease screening or diagnosis, it is often necessary to account for factors associated with the marker, such as subject characteristics or attributes of the test operator or test procedure. Analogous to approaches taken in therapeutic and etiologic studies, we adjust for covariates by characterizing the performance of the marker among subjects with fixed covariate value. We demonstrate that covariate adjustment is fundamentally different from other roles for covariates in marker evaluation. We propose the adjusted ROC curve (AROC), which is the common covariate-specific ROC curve, when covariates do not affect discrimination, and a weighted average of covariate-specific ROC curves more generally. The AROC is used to describe the age-adjusted discriminatory accuracy of prostate-specific antigen as a marker for prostate cancer in the Physicians' Health Study.

Modeling of Prostate Cancer Detection Probability with Applications

❖Robert Serfling, University of Texas at Dallas, Richardson, TX 75083, *serfling@utdallas.edu*

Key Words: prostate cancer, detection probability, finasteride

Recent results are presented on mathematical modeling of prostate cancer detection probability as a function of prostate gland volume, tumor volume, and numbers of biopsy cores in the peripheral and transition zones of the prostate. One application is to determine the number of cores needed to attain a given detection probability for a tumor of given size and a patient with given prostate gland size. Another application concerns re-interpretation of recent Prostate Cancer Prevention Trial data. For the finasteride group a reduced overall rate of tumors was observed along with, however, a higher rate of high grade tumors. The model, with the property that finasteride reduces prostate volume, as in treating benign prostate hyperplasia, supports that finasteride decreases the risk of not only low grade but also high grade tumors. Extension to model total cores "hit" is also discussed.

Empirical Likelihood-Based Inference for the Partial Area Under the ROC Curve

❖Gengsheng Qin, Georgia State University, 30 Pryor Street, Atlanta, GA 30303, *gqin@gsu.edu*; Xiaoping Jin, Centers for Disease Control and Prevention; Xiao-Hua (Andrew) Zhou, University of Washington

Key Words: ROC, AUC, The partial AUC, Diagnostic test, Confidence interval

Accurate diagnosis of disease is a critical part of health care. New diagnostic and screening tests must be evaluated based on their abilities to discriminate diseased conditions from non-diseased conditions. For a continuous-scale diagnostic test, the most commonly used global summary index of the receiver operating characteristic curve is the area under the curve (AUC). The partial area under the ROC curve (pAUC) is often used when only a region of the ROC curve is of interest. In this paper we propose an empirical likelihood approach for the inference on the partial AUC. We also conduct a simulation study to compare the relative performance of the proposed empirical likelihood based intervals with the existing normal approximation based intervals for the partial AUC. The simulation study indicates that the empirical likelihood based method greatly outperforms the existing method.

An ROC Confidence Region Based on Radial Sweep Methods

❖Michael Schuckers, St. Lawrence University, 211 Valentine Hall, Canton, NY 13617, *schuckers@stlawu.edu*; Yordan Minev, St. Lawrence University

Key Words: false accept, false reject, polar coordinates

In this paper, we present a methodology for creating a confidence region for a receiver operating characteristic (ROC) curve. This ROC curve is a measure of the ability of a classifier to distinguish between two groups. In our motivating example, we will refer to these groups as genuines and imposters. We first convert our ROC to polar coordinates for some center (c_x,c_y), a set of radii, R, and a set of angles, Q. We then bootstrap both the collection of imposters and the collection of genuines a large number of times. For each genuine and imposter replicate we calculate a new ROC, R*, in polar coordinates for the same set of angles, Q. Having done this we can find an ROC region that captures a specified proportion of the bootstrapped curves. We apply this methodology to data for classification of biometric identification systems, e.g. fingerprint recognition.

505 Inferences for Stochastic Process and Time-to-Event Data Analysis

Biometrics Section

Thursday, August 2, 8:30 am–10:20 am

Hierarchical Spatiotemporal Models for Preterm Birth

❖Jonathan Norton, Florida State University, Department of Statistics, Tallahassee, FL 32306-4330, *jondnorton@gmail.com*

Key Words: Spatial statistics, Epidemiology, Intrinsic autoregressive model, Hierarchical, Disease mapping, Spatiotemporal models

Preterm birth rates are increasing nationally. County rates show residual spatial correlation, even after adjustment for predictors such as race and poverty. The proposed model accounts for spatial dependence with an intrinsic autoregressive (IAR) prior on the risk of preterm birth, and then assumes that preterm births in a county follow a binomial distribution. Time is included to account for changes in risk factors and in the baseline risk surface. Known predictors are included in the model, and their coefficients are allowed to vary between time periods. An alternative model includes temporal autocorrelation. Methods for assessing changes in coefficients over time are considered. The models are fit to Arkansas preterm birth counts from 1994-2005. It is found that the IAR model is superior to one that does not include spatial correlation, and the apparent race effect changes over time.

Coalescent Theory for a Monoecious Random Mating Population with a Varying Size

❖Edward Pollak, Iowa State University, Department of Statistics, Ames, IA 50011-1210, *pllk@iastate.edu*

Key Words: Coalescent Theory, Monoecious Population, Varying Size

Consider a random mating monoecious population whose size varies according to an irreducible a periodic Markov chain having states 1N,...,xKN, where N is large. Its transition probabilities and stationary distribution are respectively pij=P[Ns-1=xjN|Ns=xiN],s=0,-1,-2,... and v1,v2,..,vK.. At time 0 a sample of n< < N genes is taken. Let Gij be the random number of successful gametes produced by any particular parent if Ns=xiN and Ns-1=xjN. Then coalescent theory holds, with the effective population size being the reciprocal of the sum of vipij{xj/[4N(xi squared)]}E[Gij(Gij-1)], provided that the Gijs are exchangeable, E[Gij(Gij-1)] is finite and E[Gij cubed]/N approaches 0 as N increases.

Real-Time Modeling of Infectious Disease Outbreaks

❖Laura White, Boston University, 715 Albany St T416E, Boston, MA 02118, *lfwhite@bu.edu*; Marcello Pagano, Harvard University

Key Words: Basic Reproductive Number, Serial Interval, Infectious Disease Epidemics, Disease Modeling

Infectious disease modeling is a complex problem with a long history. However, methods are lacking that adequately estimate the serial interval, defined as the time between a case and its infectee presenting with symptoms. Estimation of this interval and the basic reproductive number, R0,while an epidemic is in progress is a major public health interest that has only been addressed in a limited way. We present a likelihood-based method for simultaneous estimation of R0 and the serial interval and show how this method can be extended to adjust for a shrinking susceptible population and other population dynamics, such as the implementation of control measures. We discuss how this method can be viewed as an extension of branching process methodology, without required prior information on the serial interval. Results for this method are shown with real data.

Estimation Methods for Successive Times-to-Event Adjusted by Multiple Health-Related Quality-of-Life Scores

❖Adin-Cristian Andrei, University of Wisconsin-Madison, K6/428 Clinical Science Center, 600 Highland Ave, Madison, WI 53792-4675, *andrei@biostat.wisc.edu*

Key Words: Health-related quality-of-life, successive events, clinical trials, nonparametric estimation, multidimensional, quality-of-life-adjusted gap times

When analyzing successive or recurrent times-to-event, close attention is oftentimes paid to patient health-related quality-of-life (HRQOL). Assessing patient health status is a highly complex process, requiring that numerous HRQOL scores be collected and analyzed. HRQOL-adjusted times-to-event are useful and simple ways to summarize such complex data structures. As an application, treatment strategies that favor improved survival adjusted for one or more, possibly patient-selected, HRQOL scores, might be envisioned. Therefore, one needs to understand the joint distribution of several times-to-event adjusted by multiple HRQOL scores. Hence, using inverse probability-of-censoring weighting techniques, we propose nonparametric estimators for such joint distributions. Their performance is evaluated in simulations and their practical usefulness is illustrated by a clinical trial application.

Parameter Calibration in an Agent-Based Model of Leishmania Major Infection

❖Garrett Dancik, Iowa State University, 819 24th St, Apt 12, Ames, IA 50010, *gdancik@iastate.edu*; Karin Dorman, Iowa State University

Key Words: computer experiments, Gaussian process, parameter calibration

Computer models of disease provide in silico environments for understanding complex interactions between pathogen and host. Often, modelers take ad-hoc approaches to parameter estimation when models are high dimensional and resource intensive, so in silico results have limited real-world applicability. However, statistical methods for calibration of complex, resource-hungry models to data are available. We use a Gaussian process approximation of the computer code and estimate five immunological and pathogen-related parameters in an agent-based model of Leishmania major infection. We verify the method using simulated field data and then calibrate using published biological observations. Our results suggest that L. major has a slow growth rate and replicates for an extended time before damaging the host cell. We also discuss sensitivity analysis as a strategy for future data collection.

A Color Image Grand Tour for Exploring Medical Images

❖Juergen Symanzik, Utah State University, Department of Mathematics and Statistics, 3900 Old Main Hill, Logan, UT 84322-3900, *symanzik@math.usu.edu*; Bradley Wallet, Automated Decisions, LLC; William Shannon, Washington University in St. Louis

Key Words: IGT, Magnetic Resonance Imaging, MRI, Multispectral Images, Visualization

The image grand tour (IGT) is a method for visualizing multispectral images or multiple registered images. Past versions of the IGT projected a linear combination of the pixel vectors into one-dimensional space (for each pixel) and rendered these projected values as a gray-scale image. We will present a color version of the IGT that is better capable of highlighting hidden features than the grayscale version. Additional software enhancements such as the capability of zooming into a region of interest (e.g., a tumor) will be presented. Our main application are medical images such as multiple slices obtained from magnetic resonance imaging (MRI) scans. The IGT can be used to highlight features in MRI scans taken by different scanning techniques. It is also capable of providing a movie-like movement of soft tissues in case the MRI scans were taken at different points in time.

Least Squares Estimation in Linear Regression with Doubly Truncated Data

❖Xiaodong Luo, Mount Sinai School of Medicine, 130 West Kingsbridge Road, Department of Psychiatry, Room 2F10, Bronx, NY 10468, *xl2013@columbia.edu*; Wei-Yann Tsai, Columbia University

Key Words: Estimating function, Least squares estimator, Double Truncation

We investigate a least squares estimator for regression coefficients of the covariates in the multiple linear regression model with doubly truncated data, where the dependent variable is truncated from above and below. This estimator is a generalization of the estimator for one-sided truncated linear regression proposed by Tsai, Liu and Luo (2006), and is proved to be consistent and have an asymptotic normal distribution. The estimation procedure does not require to calculate the nonparametric estimate of the error distribution. We also discuss the construction of the confidence regions for the regression parameters. A simulation study shows that the estimator performs well with a moderate sample size.

506 Statistical Issues in Microarray Data in Medicine Discovery ● ✪

Biopharmaceutical Section

Thursday, August 2, 8:30 am–10:20 am

Using Information Theory for the Classification of Trends in Dose-Response Microarray Experiments

❖Luc Bijnens, Johnson & Johnson PRD, Turnhoutseweg 30, Beerse, 2340 Belgium, *lbijnens@prdbe.jnj.com*; Dan Lin, Hasselt University; Ziv Shkedy, Hasselt University; An De Bondt, Janssen Pharmaceutica - JJPRD; Ilse Van Den Wyngaert, Janssen Pharmaceutica - JJPRD; Tamara Geerts, Janssen Pharmaceutica - JJPRD; Timothy Perera, Janssen Pharmaceutica - JJPRD; Hinrich Goehlmann, Johnson & Johnson PRD

Key Words: microarrays, monotonic regression, information theory, minimum effective dose, dose-responses

Microarray experiments to investigate dose-responses consist of correlating expression levels of thousands of genes with several doses of treatments. Recently we developed several testing procedures to test for monotonic trends based on isotonic regressions of the observed means by using the methods of Barlow et al.(1972). Once a monotonic relationship between the gene expression and dose is established, there is a set of several possible monotone models which can be fitted to the data. A selection of the best model from this set allows us to identify both the shape of dose-response curve and the minimum effective dose level. The posterior probability of the model and the evidence ratio are calculated using information criteria which take into account both the goodness-of-fit and the complexity of the models. The method is applied to an experiment with 12 samples measuring 16998 genes.

Using Biological Information To Increase the Sensitivity of Tests for Differential Expression in Low-Dose Ionizing Radiation Studies

❖ Ruixiao Lu, University of California, Davis, 404 Russell Park, Apt 3, Davis, CA 95616, *rlu@ucdavis.edu*; Susanne R. Berglund, University of California, Davis; Alison Santana, University of California, Davis; Zelanna Goldberg, University of California, Davis; David Rocke, University of California, Davis

Key Words: microarray, gene expression, mRNA, gene group, pathway, Gene Ontology

Identification of differential expression of transcripts in human is complicated, partially because of differences in timing of transcriptional cascades between samples. We develop a new method to handle this by utilizing biological information in the analysis, after preprocessing of raw signals and analysis of the significant changes in single genes. First, select one or more pathways or genes groups and identify the probes on an array that correspond to the pathway or gene group. Second, for each probe, conduct a usual statistical analysis such as t-test for the difference between the treatment and control. We then test for whether in the aggregate this shows up or down-regulation. The significance of the test can be accessed by sampling genes or arrays. We illustrate the method with results from dose-response and time-course studies of low-dose radiation exposure in-vivo in humans.

Gene Selection: Filtering Microarray Probe Sets Based on Probe Level Consistency

❖ Willem Talloen, Johnson & Johnson PRD, Turnhoutseweg 30, Beerse, 2340 Belgium, *wtalloen@prdbe.jnj.com*; Hinrich Goehlmann, Johnson & Johnson PRD; Dhammika Amaratunga, Johnson & Johnson PRD; Luc Bijnens, Johnson & Johnson PRD

Key Words: microarray, high-dimensional data, gene selection

DNA microarray technology measures hybridization of nucleic acids to complementary molecules attached to a solid surface, referred to as probes. This high-content genomic tool typically generates many measurements of which only a relatively small subset is informative for the interpretation of the experiment. Here, we show how a probe set can be called informative or non-informative using probe level information. The underlying idea is that probes of an informative probe set detect the mRNA transcript of this probe set rather consistently if its mRNA concentration varies across arrays. This gene selection approach outperforms currently available methods, and is entirely objective. Consequently, by excluding many false positives, it offers a key solution to the main problems in the analysis of high-dimensional microarray data, being multiple testing and overfitting.

An Improved Classification Procedure for High-Dimension and Highly Noisy Data Using Data Defined Distance and Comparison Test

❖ Shuang Liu, University of California, Davis, 404 Russell Park Apt 3, Department of Statistics, Davis, CA 95616, *shnliu@ucdavis.edu*; David Rocke, University of California, Davis

Due to the high dimensionality and existence of significant noises in microarray datasets, the differences between gene groups are not well explained by many classification approaches. In this study of differentially regulated gene groups and pathways, five classification methods were first applied to each two-group classification individually. Misclassification rates were then estimated using cross-validation. Finally, a permutation test was used to improve the evaluation of the classification result by comparing the misclassification rates. By this means, we were able to reveal the latent differences between gene groups. The permutation test also showed that the Mahalanobis distance classifier out-performed the other classification methods, and the K-nearest neighbor method based on the correlation distance appeared to be a better solution to the problem of high-dimensionality.

The Trade-Off of Bias and Variability in the Analysis of Affymetrix GeneChip Probe Data

❖ Shibing Deng, Pfizer Inc., 6 Connshire Dr, Waterford, CT 06385, *shibing.deng@pfizer.com*

Key Words: microarray, bioinformatics, Genechip

Affymetrix GeneChipÆ arrays are widely used as a powerful tool to measure expression levels of thousands of genes simultaneously, particularly to measure relative expression levels between different experiment conditions. They usually require preprocesses of probe level data before analysis at gene level. These preprocesses include background correction, normalization and data transformation, etc. Their objectives are to remove systematic biases and reduce noises in the data. However, it may not always be possible to achieve the two goals at the same time. In this presentation, we demonstrate the trade-off between these two goals by applying different levels of biases in the data preprocess. We find we can significantly reduce noises in the data by including certain levels of biases, and this also improves the sensitivity and specificity of detecting the differentiate gene expression.

A Comprehensive Mixed-Model Approach for the Analysis of Microarray Gene Expression Data

❖ Jane Chang, Bowling Green State University, Department of ASOR, Bowling Green, OH 43403, *changj@cba.bgsu.edu*; Jason C. Hsu, The Ohio State University

Key Words: Optimal Design, Mixed Model, Generalized Least Squares Estimators, Variance Components, Microarray Gene Expression Data

There are many sources of systematic variation in microarray experiments which affect the gene expression levels, such as between subjects' variability (biological samples) and within subjects' sample variability (technical samples). To develop a rigorous classification algorithm and determine the required sample size to achieve pre-specified sensitivity and specificity, estimate all variance components is a necessity. In this talk, key components on how to design microarray experiment so that the all sources of variability can be estimated will be presented. A comprehensive mixed model will be used to analyze microarray gene expression data. Optimal designs for this comprehensive mixed model and its corresponding Generalized Least Squares estimators for the differences between treatments will be also presented.

Basal Gene Expression for Affymetrix Microarray and Its Application in Cancer Microarray Data Mining

❖Hui-Rong Qian, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *qianhu@lilly.com*

Key Words: basal gene expression, microarray, Affymetrix, data mining

Microarray measures the expression abundance of thousands of genes simultaneously. While in a microarray study the primary goal is usually to identify differentially expressed genes upon biological treatments, basal gene expression, or whether a gene is expressed and how strong a signal is, is equally important but less well studied. Affymetrix GeneChip is one of the widely used microarray platforms. The assessment of basal gene expression in GeneChips is perplexed by the application of perfect match and mismatch oligos and the use of multiple oligo pairs to represent a gene. We propose a method utilizing the well-established mixed model to assess the basal gene expression in Affymetrix microarray. We apply this approach in a public cancer microarray dataset and identify genes of biological interest in cancer research. We also discuss related issues in microarray data mining.

507 Business Statistics and Forecasting ● ✪

Business and Economics Statistics Section

Thursday, August 2, 8:30 am–10:30 am

Mining Spatial Patterns of Best-Performing Branches in Banking Branch Marketing Analysis

❖Wen Jun Yin, IBM China Research Laboratory, Building 19 Zhongguancun Software Park, 8 Dongbeiwang WestRoad Haidian District, Beijing, 100094 China, *yinwenj@cn.ibm.com*; Ming Xie, IBM China Research Laboratory; Jin Dong, IBM China Research Laboratory; Bin Zhang, IBM China Research Laboratory

Key Words: Spatial Data Mining, Banking Branch Analysis, Geographic Information System, Data Envelopment Analysis

Banking branch performance evaluation plays a key role in reinvigorating branch networks in competitive marketplace. As a crucial step, the successful branch patterns need to be discovered from diverse perspectives such as capacity, commercial and operational manners, geographic and demographic characteristics. However, current research and industrial efforts have mainly focused on internal patterns. In this paper, we adopt the spatial data mining technique to discover useful patterns of best branches. Data Envelopment Analysis is firstly carried out to define the best branches. Spatial data mining techniques such as clustering analysis are then employed to find out their environmental characteristics. In one real case, those successful patterns incorporating both geographic and demographic characteristics are discovered with high potential to guide real practices.

Comparing Methods Used for Handling Negative Values in Statistical Samples of Business Records

❖Eric Falk, Ernst & Young, 1225 Connecticut Ave NW, Washington, DC 20036, *eric.falk@ey.com*

Key Words: Statistical Sampling, Unmatched Credits, Negatives, Estimation, Simulation

Negative monetary values occur in business records for several reasons and must be handled appropriately to prevent biasing estimates. Different taxing authorities require different treatment of negative values in the context of statistical sampling of tax records. Typically business expense records contain debits for purchases and expenses, and offsetting credits if purchases are returned or expenses are reversed. Many of the credits can be matched to the original debit when the sampling population is being cleaned, but usually the final frame will contain unmatched credits for various reasons. This study compares four common methods currently used or required by either the IRS or state taxing authorities for handling these remaining negative values when using statistical sampling.

Dynamic Business Model Analysis for Internet-Delivered Business Services

❖Bonnie Ray, IBM Research, 43 Gesner Avenue, Nyack, NY 10960, *bonnier@us.ibm.com*

Key Words: B2B e-commerce, busines model analysis, customer segmentation, pricing optimization, marketing

The rapid increase in sales and delivery of business-oriented IT services over the Web introduces the opportunity for collecting large amounts of information not only about customer behavior in response to marketing and pricing strategies, but also about how customers configure and use the services. In combination, these data provide a rich basis for dynamically evaluating and refining policies of the underlying business model so as to achieve improved profitability. In this talk, we present analysis of data collected from small businesses using basic ERP functions, such as inventory management, sales processing, etc., over the network. In particular, we derive customer segmentation models based on customer demographics, application usage, and web site behavior and show how characteristics of these segments impact overall profitability under various business model scenarios.

Evaluating Forecasts

❖Wanli Min, IBM T.J. Watson Research Center, 1101 Kitchawan Road, Route 134, Yorktown Heights, NY 10598, *wanlimin@us.ibm.com*; Chaitra Nagaraja, University of Pennsylvania

Key Words: forecasting, model averaging, boosting

Forecasting problems arise in every aspect of business from predicting stock prices to product revenue. Innumerable approaches can be taken to model any specific problem. As business environments are continuously changing, there is often no paramount model. In such a case, it is valuable to generate several models and average over them. While this increases the stability, and hopefully accuracy, of the estimates, it also increases the complexity of the model. The latter feature makes evaluating the confidence of a forecast difficult. We propose a method to solve this issue using boosting procedures.

A Mixture of Binary Logistic Regression Analysis on Labor Supply of Mothers with Disabled Children

❖Zeng-Hua Lu, University of South Australia, School of Mathematics and Statistics, City West, GPO Box 2471, Adelaide, 5001 Australia, *zen.lu@unisa.edu.au*

Key Words: Mixture regression models, Switching regression models, Model selection criterion, EM algorithm, binary data

This paper formulates a mixture model for modeling unobserved heterogeneity of explanatory mechanism. Our model allows for different sets of regressors, and/or different interactions among the same regressors in different regression regimes. The model is demonstrated with particular interest to the binary dependent variable. A two-step procedure is proposed for model identification. The first step is to identify the number of regression regimes with each regime including all regressors. The second step is to select regressors in the regression regimes. The results of our simulation studies suggest that the procedure works well. The model is applied to study labor supply of mothers with disabled children in Australia.

A Semiparametric Approach To Estimate the Disagreement Prices Between the Hospitals and Insurers

❖Sreenivas Konda, Temple University, 203C Speakman Hall, 1810 North 13th Street, Philadelphia, PA 19122, *konda@temple.edu*

Key Words: Lognormal, Extreme order statistics, Surgery prices

The asymptotic distribution of the extreme order statistics of lognormal distribution is well known Gumbel distribution. However it is anything like Gumbel in the large enough sample case because of the slow convergence rate. In this paper theoretical and empirical evidence is provided for the slow convergence rate and recommendations are made to improve this rate. The findings of this research are applied to estimate the extremal prices of a narrowly defined surgical procedure: the minimum price a hospital is willing to accept and the maximum price an insurance provider is willing to pay. These extremal prices are then exploited to understand the bargaining power of hospitals in some metropolitan regions with different insurance providers.

Reflection on Teaching Business Statistics

❖Mammo Woldie, Texas Southern University, 3100 Cleburne, Houston, TX 77004, *woldie_mx@tsu.edu*

Key Words: Business statistics, core course, resources, quantitative competency

Business statistics is one of several courses students in business schools are required to take. Despite its applications in the different disciplines, the number of business statistics courses offered is minimal. This seems to hinder many students from appreciating the role of statistics in the different courses and in the real world. Furthermore, adequate resources need to be provided to improve the quantitative competency of the students.

508 Stochastic Processes and Related Topics ✪

IMS

Thursday, August 2, 8:30 am–10:20 am

An Extension of Fisher's Discriminant Analysis for Stochastic Processes

❖Hyejin Shin, Auburn University, 221 Parker Hall, Dept of Math and Stat, Auburn, AL 36849, *hjshin@auburn.edu*

Key Words: Fisher's discriminant analysis, reproducing kernel Hilbert space, stochastic processes

Modern data collection methods are now frequently returning observations that should be viewed as the result of digitized recording or sampling from stochastic processes rather than vectors of finite length. Our focus in this talk is on discrimination and classification in the infinite dimensional setting. Specially, we have developed a theoretical framework for Fisher's linear discriminant analysis of sample paths from stochastic processes through use of the Loeve-Parzen isomorphism that connects a second order process to the reproducing kernel Hilbert space generated by its covariance kernel. This approach provides a seamless transition between finite and infinite dimensional settings and lends itself well to computation via smoothing and regularization.

Uniform Central Limit Theorems for Density Estimators

❖Richard Nickl, University of Connecticut, Department of Mathematics, Storrs, CT 06269, *nickl@math.uconn.edu*

Key Words: density estimation, empirical processes, uniform central limit theorems, nonparametric maximum likelihood

Let p_n be some nonparametric density estimator based on n independent observations each distributed according to the law P. Define the stochastic process $f\mapsto \sqrt{n}\int f(p_n(x)dx-dP(x)$ where f ranges over some class of functions F. We say that p_n satisfies a uniform central limit theorem if the process defined above converges in law in the Banach space of bounded functions on F to a mean zero Gaussian process. We discuss recent results of this kind for classical density estimators such as trigonometric series, (sieved) maximum likelihood and kernel density estimators. (Part of this is joint work with Evarist Gine).

Detection of Two-Sided Alternatives in a Brownian Motion Model

❖Olympia Hadjiliadis, Princeton University/Drexel University, 8234 Taylor Ct, Princeton Junction, NJ 08550, *ohadjili@princeton.edu*; Vincent Poor, Princeton University

Key Words: Change-point detection, Brownian motion, 2-CUSUM

This problem is about detecting a change in the drift of a Brownian motion. Traditionally, 2-CUSUM stopping rules have been used for this problem due to their asymptotically optimal character. In particular, attention has focused on 2-CUSUM harmonic mean rules due to the simplicity in calculating their first moments. In this paper, we derive closed-form expressions for the first moment of a general 2-CUSUM stopping rule. We use these expressions to obtain explicit upper and lower bounds for it. Moreover, we derive an expression for the rate of change of this first moment as one of the threshold parameters changes. Based on these expressions we obtain explicit upper and lower bounds to this rate of change. Using these expressions we are able to find the best 2-CUSUM stopping rule with respect to an extended Lorden criterion both in the case of symmetric and nonsymmetric change.

Spectral Approaches for Clustering Time Series with Long Memory and Short Memory

❖Jaydip Mukhopadhyay, University of Connecticut, 215 Glenbrook Rd U4120, Storrs, CT 06269, *jaydipm@stat.uconn.edu*; Nalini Ravishanker, University of Connecticut; Cyr E. M'lan, University of Connecticut; Jonathan Hosking, IBM T. J. Watson Research Center

Key Words: periodogram, distance matrix, similarity

We discuss the problem of comparisons of a large number of time series exhibiting long memory and short memory behavior. Assessment of similarity in such time series is useful in applications to environmental, financial, or network traffic data. A hierarchical algorithm enables effective clustering using a quasi-distance matrix constructed from second order spectral properties of the time series.

Estimation of Age of a Galton Watson Process

❖Satrajit Roychoudhury, Schering Plough Research Institute, 2015 Galloping Hill Road, Building K 15, Kenilworth, NJ 07033, *satri_ray@hotmail.com*; Pranab K. Sen, The University of North Carolina at Chapel Hill; Manish C. Bhattacharjee, New Jersey Institute of Technology

Key Words: Braching Process, Galton Watson Process, EM Algorithm, Age

There are several situations where we might want to estimate the age of a Markov process. For example, we may be interested in estimating the length of time a species has been in existence in its present form, without having much historic information such as fossils for carbon dating. Also, in the genetic context, the usual problem is to find the age of an allele, given its current frequency, which means, estimating the time point when a mutation took place. For a discrete time branching process or Galton Watson Process, its age is the generation that corresponds to the earliest observed value. Three different methods of estimation of `age' are discussed along with examples. An EM algorithm is developed to estimate the parameters of offspring distribution. All the estimators have been developed in the nonparametric setup, assuming no parametric form of the offspring distribution.

Stability of Empirical Minimizers Through Gaussian Approximation

❖Philippe Berthet, The University of Rennes 1, IRMAR Universite Rennes 1, Campus de Beaulieu, Rennes, 35700 France, *philippe.berthet@univ-rennes1.fr*

Key Words: Empirical minimization, Stability properties, Strong invariance principles, Empirical processes, Donsker classes

In recent works with D.M. Mason we have developed strong invariance principles which allow one to take advantage of the nice Gaussian nature of empirical processes. Let consider the usual setting in functional data analysis, non-parametric regression, high dimensional classification, model selection and learning theory : an empirical risk minimizer f_n is selected among a class F of functions, large compared to the size n of the available i.i.d. sample with law P. Thanks to the Gaussian approximation we evaluate the true L_1 or L_2 volume of a r-neighborhood N of f_n in terms of F and the deterministic empirical L_1 radius r. The ERM algorithm is stable if this volume is small with high probability. This is the case when r is small enough compared to a quantity depending on F, P and square root of n. We also study the small probability of catching the true minimizer in N.

509 Bayesian Methods

Section on Bayesian Statistical Science

Thursday, August 2, 8:30 am–10:20 am

Bayesian Inference Resistant to Outliers Using Super Heavy-Tailed Distributions for the Calculation of Premiums

❖Alain Desgagne, University of QuÈbec at MontrÈal, CP 8888 Succ Centreville, Montreal, QC H3C 3P8 Canada, *desgagne.alain@uqam.ca*; Jean-FranÁois Angers, University of MontrÈal

Key Words: Bayesian Inference, Outlier, Heavy-Tailed Modeling, Scale Parameter, Actuarial Science, Insurance Premiums

We assume that the claims for several risks are conditionally independent random variables given the same scale parameter. The Bayesian predictive distribution of a next claim size is used to estimate the pure premium. Robust procedures to conflicting information (prior or outliers) depend mainly on the tail behavior of the likelihood and prior densities. Simple conditions are established to determine the proportion of observations that can be rejected as outliers. It is shown that the posterior distribution converges in law to the posterior that would be obtained from the reduced sample, excluding the outliers, as they tend to 0 or infinity, at any given rate. We compare the log-normal model with the robust super heavy-tailed (log-Pareto type) distributions model. An example of calculation of a pure premium is given.

Redefining Degrees of Freedom for General Hierarchical Linear Models

❖Yue Cui, The University of Minnesota, 1119 E River PKWY, Minneapolis, MN 55414, *yuecui@biostat.umn.edu*; James S. Hodges, The University of Minnesota

Key Words: degrees of freedom, hierarchical models

Degrees of freedom (DF) have been used in frequentist methods for significance tests. The idea has been extended to Bayesian methods to measure complexity of hierarchical models. Spiegelhalter et al. (2002) proposed the "effective number of parameters," p_D, which is easily computed but has practical problems (Lu et al. 2006). Hodges and Sargent (2001) extended linear model theory to give another complexity measure (DF) for a hierarchical model's overall fit. In special cases, H&S's DF factors naturally into effect-specific DF, but in general, it is unclear how to attribute DF to individual effects. We give a new definition of DF that naturally decomposes a fit's total DF into effect-specific DF for arbitrary normal-error linear hierarchical models. This gives a way to place DF-based priors on smoothing parameters, and to describe the complexity of individual effects.

Empirical Bayes Models of Poisson Clinical Trials

❖Boris Zaslavsky, Food and Drug Administration, 1401 Rockville Pike, CBER, Rockville, MD 20852, *boris.zaslavsky@fda.hhs.gov*

Key Words: gamma distribution, negative - binomial distribution, maximum likelihood, Poisson distribution, WinBUGS

The Gamma distribution is a natural prior for Poisson models. Under the empirical Bayes approach, the parameters of this distribution are the maximum likelihood estimator of the marginal negative-binomial distribution. The straightforward numerical search for the maximum likelihood solution is impractical given available software. We propose a simplification to the maximum likelihood problem, by preliminary implementing Markov Chain Monte Carlo (MCMC) method using software WinBUGS. The results of WinBUGS calculations are used as a starting point in the maximum likelihood estimation of gamma distribution. Easily computable approximation formulae may be used to find the maximum likelihood point estimations for the gamma distribution. The solution for this approximation converges to the precise maximum likelihood solution as the sample size increases.

A New Approach to the Inverse Bayes Formula of Compatible Conditional Distributions

❖Kun-Lin Kuo, National Chengchi University, Dept. of Math, 64 ChiNan Road Section 2, Wen-Shan, Taipei, 11605 Taiwan, *g1751501@nccu.edu.tw*; Chwan-Chin Song, National Chengchi University; Thomas J. Jiang, National Chengchi University

Key Words: compatibility, joint density, conditional density, marginal density, rank one positive extension

The inverse Bayes formula (IBF) is a statistical tool which is important in distribution theory and Bayesian missing data problems (examples can be seen in Tian and Tan (2003), among others). Ng (1997) provided a form of IBF essentially for product measurable space and recognized that IBF is potentially useful in computing marginals and checking compatibility. Recently, Tian and Tan (2003) provided a form of modified IBF in nonproduct measurable space and gave some applications. Although the modified IBF is useful, it holds only in a restricted space. In this article, we give a new theoretical approach to the IBF that can be applied in more general space and in both discrete and continuous cases. This new approach provides the rationale of Ng (1997)'s IBF and may also solve some problems that Tian and Tan (2003) can't do. Examples are given to illustrate our new approach to the IBF.

Bayes Sequential Estimation for a Subclass of Exponential Family of Distributions Under LINEX Loss

❖Alicja Jokiel-Rokita, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, Wroclaw, 50-370 Poland, *alicja.jokiel-rokita@pwr.wroc.pl*

Key Words: sequential procedure, Bayesian estimation, exponential family, APO rule, AO rule, LINEX loss function

A subclass of an exponential family of distributions is considered. The problem is to estimate an unknown distribution parameter of this family. As an estimation error a LINEX loss function is used and a constant cost for each observation is assumed. The Bayesian approach to the problem is applied assuming conjugate prior distributions of the unknown parameter. A class of asymptotically pointwise optimal (APO) and asymptotically optimal (AO) procedures is derived.

Estimation Procedures in the Case of Randomly Forthcoming Data

❖Ryszard Magiera, Wroclaw University of Technology, Wybrzeze Wyspianskiego 27, Wroclaw, 50-370 Poland, *ryszard.magiera@pwr.wroc.pl*; Alicja Jokiel-Rokita, Wroclaw University of Technology

Key Words: sequential procedure, Bayesian estimation, stopping time, exponential family

The following statistical model is considered. The collection of data from several independent populations is available only at random times determined by order statistics of lifetimes of a given number of objects. Each of the populations is distributed according to a general multiparameter exponential family. The problem is to estimate the mean value vector parameter of the multiparameter exponential family of distributions of the forthcoming observations. Under the loss function involving a weighted squared error loss, the cost proportional to the events appeared and a cost of observing the process, a class of optimal sequential procedures is established. The procedures are derived in two situations: when the distribution of the lifetimes is completely known and in the case when it is unknown but assumed to belong to an exponential subfamily with an unknown parameter.

Bayesian Analysis of Progressively Type-II Censored Sample from Burr Type-XII Distribution

❖Younshik Chung, Pusan National University, 30 Jangjun Dong Guemjung Gu, Department of Statistics, Busan, Republic of Republic of Korea, *yschung@pusan.ac.kr*; Seongho Song, University of Cincinnati; Chansoo Kim, Kongju National University; Junghoon Jang, Korea Food and Drug Association

Key Words: Progressively Type-II Censored sample, Burr Type-XII distribution, Expected Fisher Information, Bayesian Estimation, Bayesian Prediction

In this article, the estimation of parameters based on a progressively Type-II censored sample from a Burr Type-XII distribution is investigated by using maximum likelihood (ML) approach as well as Bayesian approach. The exact expression of the expected Fisher information matrix of parameters in the distribution has been obtained. Based on this result, we construct the confidence interval for the parameters using the asymptotic normality of the ML estimator. We compare it to Bayesian credible interval. We also consider the problem of predicting an independent future observation from the same distribution. The Gibbs and Metropolis samplers are used to predict the behavior of future observations from the same distribution.

510 Modeling and Simulation

Section on Statistical Computing, Section on Statistical Graphics

Thursday, August 2, 8:30 am–10:20 am

Confirmatory Factor Analysis and Measurement Error: A Monte Carlo Simulation

❖Bruce Brown, Brigham Young University, 1098 SWKT, Provo, UT 84602, *bruce_brown@byu.edu*; Tyson T. Phalp, Myriad Genetics; Petr Blahus, Charles University

Key Words: confirmatory factor analysis, structural validity, monte carlo simulation

It has been demonstrated by Lauritzen et al (2007) that exploratory factor analysis cannot produce structurally valid results unless a minimum reliability of r = .50 is met, and very few factor analyses in published studies meet these reliability requirements. The present study examines the operating characteristics of confirmatory factor analyses using the same Monte Carlo simulation methods. A variety of "manifest data" configurations are created from known underlying latent variable structures. Comparisons are drawn between the structural validity results for confirmatory factor analyses and those for exploratory factor analyses. It is demonstrated that when one knows how many factors are present in the latent variable structure, confirmatory factor analysis is somewhat more effective than exploratory factor analysis in producing structurally valid results.

Maximum Likelihood–Based Confidence Intervals for Complementary Poisson Rate Parameters with Data Subject to Misclassification

❖Kent Riggs, Stephen F. Austin State University, Box 13040 SFA Station, Nacogdoches, TX 75962-3040, *riggske@sfasu.edu*

Key Words: asymptotic intervals, misclassification, double sampling, simulation

Of interest is statistical inference for two complementary Poisson rate parameters where the data is subject to misclassification. A double sampling scheme is employed to account for misclassification and to estimate the binomial misclassification parameters. We propose an EM algorithm to find restricted and unrestricted MLEs. Next, we derive the maximum-likelihood-based confidence intervals for the Poisson and binomial parameters using the Wald, score, and likelihood ratio statistics. We then study coverage and width properties of the three asymptotic intervals. Finally, we apply the confidence intervals to a real-life data set.

Fitting Curves to Nested Confidence Intervals

❖Bob Blodgett, Food and Drug Administration, 5100 Paint Branch Parkway, College Park, MD 20740, *robert.blodgett@fda.hhs.gov*

Key Words: Risk assessment

A curve can be fit to a confidence interval and an estimate it contains. This allows the uncertainty in the measurement to be used with the estimate. When nested confidence intervals are available at several different levels, they can be fit with a single curve defined piecewise.

The Semiparametric Exchangeable Model

❖Stephine Keeton, Food and Drug Administration, 7500 Standish Place, Rockville, MD 20855, *stephine.keeton@fda.hhs.gov*

Key Words: exchangeability, semi parametric

We generalize the class of exchangeable parametric distributions and present a unified model. We also propose a semi parametric regression model and develop procedures of efficient estimation. Because independence is relaxed to exchangeability, the unified model for exchangeability improves models for which independence is assumed. Semi parametric models, as a modern regression, share the advantages of both parametric and nonparametric models. The proposed semi parametric exchangeable model is anticipated to ameliorate statistical modeling, in particular, in teratology.

A Parametric Bootstrap Approach for ANOVA with Unequal Variances: Fixed and Random Models

❖Fei Lu, University of Louisiana at Lafayette, POBox 40678, Lafayette, LA 70504, *lufei@louisiana.edu*; Kalimuthu Krishnamoorthy, University of Louisiana at Lafayette ; Thomas Mathew, University of Maryland, Baltimore County

Key Words: Fixed Effects, Generalized F test, Generalized p-value, Interlaboratory studies, Random Effects, Welch test

This article is about testing the equality of several normal means when the variances are unknown and arbitrary. Even though several tests are available in the literature, none of them perform well in terms of type I error probability under various sample size and parameter combinations. We propose a parametric bootstrap (PB) approach and compare it with three existing location-scale invariant tests: the Welch test, James test, and the generalized F test. The size and power properties of the tests are evaluated using Monte Carlo simulation. Our studies show that the PB test is the best among the four tests with respect to size properties. It is also noted that the same tests can be used to test the significance of the random effect variance component in a one-way random model under unequal error variances. Such models are widely used to analyze data from inter-laboratory studies.

Some Binary and Nonbinary Construction Plans for Diallel Cross Designs in Incomplete Blocks

❖Jeffrey Shaffer, Tulane University, 5855 Walnut Creek Road, Apt b210, River Ridge, LA 70123, *jshaffer@tulane.edu*; Sudesh Srivastav, Tulane University

Key Words: block design, diallel cross, construction plans, binary design, non-binary design, mating

A general construction technique is introduced for deriving diallel cross designs using designs with nested rows and columns. Infinite series of diallel cross designs are derived for arrangements in both binary and nonbinary blocks. Efficiency implications of conducting diallel cross experiments in nonbinary blocks are discussed.

A Continuing Study on a New Resampling Method To Reduce Small-Sample Bias: An Extension to Multivariate Data

❖Haiyan Bai, University of Central Florida, College of Education, 4000 Central Florida Blvd, PO Box 161250, Orlando, FL 32816-4880, *hbai@mail.ucf.edu*; Wei Pan, University of Cincinnati; Stephen Sivo, University of Central Florida

Key Words: resampling, bootstrap, jackknife, randomization tests, cross-validation, small-sample bias

The four extant resampling methods-the bootstrap, and its counterparts: the jackknife, randomization tests, and cross-validation-are applicable to multivariate data, but with limitations of finding a single-valued statistic or combining the univariate p-values into a single p-value (Good, 2006). Recently, a new resampling method was developed to produce amplified samples with less bias, more robust to outliers, and larger statistical power than the pre-existing resampling methods; unfortunately, the new method is limited its application to univariate data. The current study extends the new method to multivariate data without the aforementioned limitations, providing a practical technique with a plug-in principle for researchers to employ advanced analysis directly on the amplified multivariate data with retained statistical properties, including correlations.

511 Useful Advice in Teaching Statistics

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Thursday, August 2, 8:30 am–10:20 am

Teaching Introductory Statistics to Prospective Elementary School Teachers: The Habitat Experience at ETSU

❖Edith Seier, East Tennessee State University, 10 Brooklawn Ct, Johnson City, TN 37604, *seier@etsu.edu*; Susan Hosler, East Tennessee State University

Key Words: Projects, posters, K-6 standards, hands-on activities, reading assignments, resources web-page

A statistics course is the mathematics general education requirement at ETSU. Recently, as part of an initiative to improve the education in science and mathematics of prospective elementary school teachers, a special section has been created for them. The standard topics of an algebra-based introductory statistics course are still covered and students take the departmental comprehensive final. In-class activities and assignments have been added to increase student's awareness about the teaching of statistics and probability in the elementary school, to familiarize them with available resources and to expose them to learning strategies and tools. We are working toward a stronger communication between statistics and science courses, projects being a bridge between them. Activities will be described and student response and ideas for improvement will be discussed.

Differences in Students' Performance Due to a Statistics Lab Component

❖Jose H. Guardiola, Texas A&M University, Corpus Christi, 6300 Ocean Drive, Unit 5825, Corpus Christi, TX 78412-5825, *jose.guardiola@tamucc.edu*; Duran-Hutchings Nadina, Texas A&M University, Corpus Christi; Joseph Champine, Texas A&M University, Corpus Christi

Key Words: students performance, statistics lab, instructional methods, elementary statistics, students achievement

In this paper we discuss the performance of students on elementary statistics courses by comparing two similar courses taught by the same instructor. One course has a lab component using scenarios and an extensive use of technology; the second course is more traditionally taught, without a lab component. Initially an aptitude test is given to students aiming to use these values as covariates. Different measures of achievement are used on both groups in order to assess students' performance. An analysis of covariance is performed in order to detect differences on achievements on these two groups. Conclusions are drawn from the comparison of these methods. This study is aiming to evaluate the convenience of including a statistics lab in an elementary statistics course and to identify student characteristics that may contribute to a superior achievement on one of these two methods.

Undergraduate Consulting as an Introductory Course

❖Jim Bentley, University of Redlands, 7499 Holloway Road, Rancho Cucamonga, CA 91730, *jim_bentley@redlands.edu*

Key Words: Undergraduate, Consulting

The University of Redlands offers an undergraduate consulting course in which students participate in the entire consulting process. The course is an upper division math class with a listed prerequisite of an introductory statistics class. In reality, except for math majors, the only prerequisite is a college level algebra class. This talk will address how projects are obtained from faculty and administrators, industry, government, and investigators of national health studies. The statistical topics to which the students are exposed will be described. Information will also be given as to how the class is structured to provide the students with an opportunity to experience consulting from the viewpoint of the statistician. The end product of the class is a group of students who are prepared to work with statisticians within their own fields or to continue their studies in statistics.

Mentoring an Undergraduate Research Project

❖Debra Hydorn, University of Mary Washington, 27 Huntington Drive, Fredericksburg, VA 22405, *dhydorn@umw.edu*

Key Words: Statistics Education, Undergraduate Research

As part of Mary Washington's Summer Science Institute, I had the opportunity to mentor two students on two research projects in 2006. Because Mary Washington does not offer a statistics degree, both students were mathematics majors and the only statistics courses they had taken before the summer were an introductory course and a one semester upper-level probability and statistics course. Neither student had any programming experience or exposure to LaTex for writing up their results. In this session I will describe the projects and computer simulations the students completed and the process I followed to prepare them for the ten week summer program.

Teaching a 'History of Statistics' Discussion Course

❖Phyllis Curtiss, Grand Valley State University, Department of Statistics, 1 Campus Drive, Allendale, MI 49401, *curtissp@gvsu.edu*; Kirk Anderson, Grand Valley State University

Key Words: History, David Salsburg, Discussion

Is it possible to combine "history" and "statistics" into an EXCITING course, not a "dull" and "boring" one, with a high level of student interest? Yes! At Grand Valley State University, we offer a one-credit reading/discussion class centered on David Salsburg's book "The Lady Tasting Tea: How Statistics Revolutionized Science in the 20th Century." Students are given a reading assignment each week which is discussed the next class period. In this talk, we will discuss how the class got started and the format we use for instruction. We will also share ideas to get the students involved and some feedback from students that have taken the class.

Writing Assignments in Elementary Statistics Classes

❖Ananda Jayawardhana, Pittsburg State University, 219 C, West Hudson, Pittsburg, KS 66762, *ananda@pittstate.edu*

Key Words: Writing to learn, Elementary Statistics

In this paper we explain the methods and writing assignments we have used in Elementary Statistics classes at Pittsburg State University. Daily journal writing, weekly papers with multiple submissions, and daily in class writing are some of the methods we have utilized. Our experience about the immediate feedback has been eye opening. Students have mixed feelings about writing questions in quizzes and exams. Student's comments about the writing to learn approach is also discussed.

A Random Walk Through ANOVA

❖Daniel Kaplan, Macalester College, Dept of Mathematics, 1600 Grand Avenue, Saint Paul, MN 55116, *kaplan@macalester.edu*

Key Words: ANOVA, education, statistical modeling, graphics, geometry

We often teach ANOVA as an algorithm for generating F statistics and p-values. I'll introduce another way, based on simple ideas of random walks, that makes it clear why we talk about mean squares. This way of presenting ANOVA makes clear the intimate connection between ANOVA and regression, and allows students to move easily beyond the limiting and incorrect notion that ANOVA is just a generalization of the t-test to more than two groups.

512 Topics in Environmetrics

Section on Statistics and the Environment, WNAR

Thursday, August 2, 8:30 am–10:20 am

Nonlinear Forecasts of the Gulf of Maine Ecosystem

❖Barbara Bailey, San Diego State University, Dept of Mathematics and Statistics, 5500 Campanile Drive, San Diego, CA 92182-7720, *babailey@sciences.sdsu.edu*; Andrew Pershing, University of Maine

Key Words: neural network models, prediction intervals, bootstrap

Time series of physical and biological properties of the ocean are a valuable resource for developing models for ecological forecasting. Both the physics of the oceans and organisms living in it can exhibit nonlinear dynamics. We describe the development of a model that forecasts the abundance of Calanus finmarchicus from hydrographic data from the Gulf of Maine. We present the results of a neural network model including model diagnostics and forecasts. Measures of uncertainty and prediction intervals are obtained by the nonparametric bootstrap.

Nonparametric Smoothing of Irregularly-Spaced Palaeo Records: Trend Estimation with Application to Climate Change

❖Patricia Menendez, Swiss Federal Institute for Forest Snow and Landscape Research WSL, Zuercherstrasse 111, Birmensdorf, CH-8903 Switzerland, *patricia.menendez@wsl.ch*; Sucharita Ghosh, Swiss Federal Institute for Forest Snow and Landscape Research WSL

Key Words: Palaeo time Series, Climate Change, Irregularly spaced data, Kernel smoothing, Long-range dependence, Trend estimation

Palaeo data such as fossil pollen or ice isotope records are irregularly spaced in time due to the nature of the dating procedure. Such studies consider the evolution of the processes along the age scale. In statistics, uneven and random sampling schemes in time series data have been considered by various authors, in particular for alias-free spectral estimation (e.g., Shapiro and Silverman 1960, Masry 1978 and others). We consider the nonparametric smoothing problem, when the data have been sampled irregularly and randomly in time. As for the model, of special interest are processes with time varying distribution functions, such as a time dependent one-dimensional transformation of a stationary Gaussian process; see Ghosh, Beran and Innes 1997 and Ghosh and Draghicescu 2002. Asymptotic results under long-range and short-range dependence are derived with examples of fossil records.

A Test of Equality of Spectra of Short Memory Independent Gaussian Time Series

❖Ravi Siddani, University of Maryland, Baltimore County, 5003 Westland Boulevard, Apt H, Baltimore, MD 21227, *siddani1@umbc.edu*

Key Words: Spectral density, Autocovariance function, Periodograms, Whittle Likelihood, Contiguity

We develop a test of equality of the spectral densities of short memory independent Gaussian time series. This problem is equivalent to the equality of autocovariance functions. The autocovariance appears nonlinearly in the normal likelihood and is difficult to tackle analytically. We propose a test based on the asymptotic independence of the periodograms, and approach our null hypothesis of equality of spectral densities by an increasing sequence of hypotheses of equality of spectra at the Fourier frequencies. We examine the performance of our test statistic as the sample size increases. Appealing to the contiguity of the Whittle likelihood (an explicit function of the spectral density) and the normal likelihood we show that our test statistic maintains the size for large sample sizes. We apply our methodology to a rain gauge data set from Melbourne, Florida.

Temporally Correlated Dirichlet Processes in Pollution Receptor Modeling

❖Matthew Heaton, Brigham Young University, 1441 N 1450 E, Provo, UT 84604, *mattheaton@gmail.com*

Key Words: Autocorrelation, Dynamic Models, Source Profiles, Air Pollution, Dirichlet Process

Understanding the effect of pollution arising from human activity on the environment is an important precursor to promoting public health and environmental sustainability. One aspect of understanding pollution is understanding pollution sources. Multivariate receptor modeling seeks to estimate pollution source profiles and pollution emissions from concentrations of pollutants such as particulate matter (PM) in the air collected over consecutive time periods. Previous approaches to multivariate receptor modeling assume independence of PM measurements and constant source profiles. Notwithstanding, the existence of temporal correlation among PM measurements is commonly accepted. In this paper an approach to multivariate receptor modeling is developed in which the temporal structure of PM measurements is accounted for by modeling source profiles as a time dependent Dirichlet process.

Estimation of Sea Depth by Data Assimilation with Tsunami Simulation Model

❖Kazuyuki Nakamura, The Graduate University for Advanced Studies, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, International 1068569 Japan, *nakakazu@ism.ac.jp*; Tomoyuki Higuchi, The Institute of Statistical Mathematics; Naoki Hirose, Kyushu University

Key Words: data assimilation, particle filter, state space model, tsunami

A tsunami simulation model can forecast the arrival time and the height of tsunamis. The sea bottom topography is included in the model as the boundary condition. However, the bottom topography datasets have errors which cause inaccurate forecasts. Therefore, we should modify the bottom topography in the model to obtain more precise results. To correct the topography, we have introduced a data assimilation framework in which a tsunami simulation model and tide gauge records are combined. We will demonstrate the framework in terms of nonlinear state space model and present the result applied to the Okushiri Tsunami, which occurred in the Japan Sea in 1993. The particle filter is used in this estimation. The result indicates that an area in the Japan Sea might be shallower than the existing topography datasets.

Bias of Cormack-Jolly-Seber (CJS) Survival Estimators in Release Recapture Studies with Low Sampling Effort

❖James R. Faulkner, National Marine Fisheries Service, 2725 Montlake Blvd East, Seattle, WA 98112-2097, *jim.faulkner@noaa.gov*; Steven G. Smith, NOAA Fisheries

Key Words: estimator bias, release-recapture, Cormack-Jolly-Seber, sampling, survival

CJS maximum likelihood (ML) estimators for survival probabilities and their associated variances are nearly asymptotically unbiased, but small effective sample sizes can result in substantial bias. Release-recapture experiments that follow a single marked release with low sampling effort are particularly susceptible to estimator bias. We use computer simulation to investigate the bias of CJS ML estimators and some bias-adjusted estimators of survival under conditions of low sampling effort. We also investigate the bias of the large sample variance of the survival estimators under various conditions.

Climate Change Impacts and Models

❖John Aleong, University of Vermont, Mathematics and Statistics, Burlington, VT 05405, *John.Aleong@uvm.edu*

Key Words: Simulation models, environment, transportation systems, infrastructure, case studies

Global climate change may be one of the main problems of the 21st century. Many simulations models are used to predict increasing global temperature, increasing carbon dioxide, shifts in weather patterns and irreversible rise in sea levels. We will review some of the scientific models, the global impact of climate changes on the environment, society, transportation systems and infrastructure, with some case studies. The latest report of the Intergovernmental Panel on Climate Change (IPCC), a U.N. body will be discussed.

513 Optimal Sample Design ●

Section on Survey Research Methods

Thursday, August 2, 8:30 am–10:20 am

Survey Designs To Optimize Efficiency and Precision for Multiple Objectives: Methods and Applications

❖Stephen Williams, Mathematica Policy Research, Inc., 148 Fletcher Ct, Cary, NC 27511, *swilliams@mathematica-mpr.com*; Frank J. Potter, Mathematica Policy Research, Inc.

Key Words: design optimization, nonlinear programming, multiple survey objectives, multi-mode surveys

Allocation of the sample among strata or sample clusters on the basis of variance components and survey costs is important to survey design. Two basic approaches to solving for optimum are: maximize precision for a fixed cost or minimize cost for a specified precision. Optimum allocation equations for an estimated mean or total of a specific population are presented in sampling methods books. However, we typically need an optimum allocation that simultaneously satisfies several types of estimates and for several inference subpopulations. In this paper we review optimization methodology, its history and its extension to such multiple survey objectives. The computer algorithm we used to solve this nonlinear equation problem is described. Two recent applications are used to demonstrate the diversity of optimization problems and the flexibility of the methodology.

Optimization of Nonproportional Samples

❖Raimund K. Wildner, GfK Group, Nordwestring 101, Nuernberg, D 90319 Germany, *raimund.wildner@gfk.com*

Key Words: Stratification, Weighting, Sampling, Optimization, Effective sample size

The well-known optimum distribution of stratified samples by Neyman-Tschuprow minimizes the sampling variance of the total mean of one variable. In practical surveys however, there are many variables that are to be reported with reasonable small variance not only for the total but for numerous segments. A solution is presented that minimizes the variance of the weighted sum of multiple means. Examples are shown that demonstrate how the solution behaves under different conditions. Furthermore, disproportional surveys have to be weighted to get unbiased results. An important parameter to evaluate the quality of a weighted sample is the effective sample size. But the well known formula for the effective sample size is valid only for proportional samples. The optimization formula can as well be used to derive a formula for the effective sample size for non proportional samples.

Sampling Design and Inference for Social Network Data: Estimating Epidemic Potential with Contact Tracing Data

❖Krista Gile, University of Washington, Box 354320, Department of Statistics, Padelford Hall C14, Seattle, WA 98195-4320, *kgile@stat.washington.edu*; Mark S. Handcock, University of Washington

Key Words: survey, social science, random graph model, epidemiology

Social network models are of increasing importance for applications from disease spread to social capital to national security. Numerous sampling strategies are used to gather network data for model fitting, including egocentric and link-tracing designs. In the case of reportable sexually transmitted diseases, network sampling is often conducted through “contact tracing,” in which only infected individuals are sampled. This talk applies principles of statistically optimal design to sampling relational data. In particular, we study various contact tracing designs used to collect data to fit exponential family random graph models and estimate epidemic potential. We report advances in sampling and optimal design relevant to social science and epidemiological research.

Sample Allocation Under a Population Model and Stratified Inclusion Probability Proportional to Size Sampling

Steven Heeringa, Institute for Social Research; ❖SunWoong Kim, Dongguk University, Jung-Gu Pil-Dong 3 Ga 26, Seoul, 100-715 Korea, *sunwk@dongguk.edu*; Peter Solenberger, Institute for Social Research

Key Words: Superpopulation Regression Model, Stratified Sampling, Sample Allocation

Many studies have been conducted on sample allocation for stratified sampling designs. We consider optimal allocation for stratified sample designs in which there is a population regression model for the variable of interest. We present a flexible approach for optimal sample allocation to minimize the expected variance of the Horvitz-Thompson estimator under the population regression model. We show that the presence of the intercept in the model produces a comparatively complicated allocation problem. For one popular IPPS sampling scheme we show how to determine stratum sample sizes based on optimization theory.

Comparison of Sample Set Curtailment Procedures

❖Jason Legg, Iowa State University, CSSM, 216 Snedecor, Ames, IA 50011, *jlegg@iastate.edu*

Key Words: Rejective sampling, Cube sampling, Stratification, Balanced sampling

For many designs, there is a nonzero probability of selecting a sample that provides poor estimates for known quantities. Stratified random sampling reduces the set of such possible samples by fixing the sample size within each stratum. However, undesirable samples are still possible with stratification. Rejective sampling removes unwanted samples by only retaining a sample if specified functions of sample estimates are within a tolerance of known values. The resulting samples are often said to be balanced on the function of the variables used in the rejection procedure. Cube sampling, an alternative to rejective sampling, attempts to select a balanced sample with the same first-order inclusion probabilities as the original design. Through simulation, we compare estimation properties of a rejective sampling procedure to those of cube sampling for estimating a mean.

A Model-Assisted Design Based on Adaptive Sampling

❖Ismaila Adeleke, University of Lagos, Department of Actuarial Science and Ins, Lagos Nigeria, 101017 Nigeria, *adeleke22000@yahoo.ca*; Raymond Okafor, University of Lagos; Ebenezer Esan, University of Lagos; Athanasius Opara, University of Lagos

Key Words: Model-based, design based, adaptive sampling, inclusion probabilities, sampling weights

Two problems often crop up in adaptive sampling. One, it may not be feasible to sample according to a designated sampling plan. And two, the designated sampling plan may result in very small selection probabilities for same units thereby giving large weights to such units in estimation. In order to ameliorate these problems, we propose a regression procedure that combines design and model-based techniques of inference.

514 Multivariate and Longitudinal Analysis of Survey Data ●

Section on Survey Research Methods

Thursday, August 2, 8:30 am–10:20 am

Random Forests with Survey Data

❖Guillermo Mendez, Arizona State University, Dept of Mathematics and Statistics, Tempe, AZ 85287-1804, *guillermo.mendez@asu.edu*; Sharon L. Lohr, Arizona State University

Key Words: random forest, importance scores, clustered data

Random forests (Breiman, 2001) are a powerful tool for modeling data with both a categorical and continuous response. Although they are used primarily for prediction, a byproduct of the random forest algorithm is estimates of the importance scores of each predictor. One assumption for random forest, that the data in the training set is independently generated, will not necessarily hold when the data are clustered. We explore methods for estimating importance scores of potential variables for data from clustered samples.

Multivariate Analysis Techniques on Model-Based Sampling

Chang Tai Chao, National Cheng Kung University; ❖Feng Min Lin, National Cheng Kung University, Department of Statistics,, No.1, Ta-Hsueh Road, Tainan, 701 Taiwan, *fmlin@stat.ncku.edu.tw*

Key Words: Model-Based Sampling, Optimal Sampling Strategy, Multivariate Analysis, Principal Component, Cluster Analysis, Canonical Analysis

Previous work regarding lowering prediction error in model-based sampling strategies have different shortcomings. For example, optimal strategies often require intensive computation, also depend on the population model and the predictor. The spatial systematic design cannot work properly under an anisotropic population, or population with nonhomogeneous variance. Intuitively, one would like to select units that account for as much population variability as possible. Ideally, it can be done by carefully examining the population covariance matrix, in which the information of the population variability is contained. The object of this study is to construct several model-based sampling designs which make use of various multivariate analysis techniques to explore the information contained in the population covariance matrix. The properties of these designs will be discussed and compared.

A Multivariate Time Series Model for Small Areas in the Dutch Labor Force Survey

❖ Sabine Krieg, Statistics Netherlands, PO Box 4481 Room R406, Heerlen, 6401CZ Netherlands, *skrg@cbs.nl*; Jan Van den Brakel, Statistics Netherlands

Key Words: structural time series models, Kalman filter, small area estimation, timeliness, unemployment

The Dutch Labor Force Survey (LFS) is based on a rotating panel design. The monthly sample size is too small to produce reliable monthly estimates for the unemployment at both the national level and domeins with the generalized regression estimator. Therefore the monthly unemployment figures are based on the data observed in the preceding three months. In this paper a multivariate structural time series model for six domains is applied to the data of the LFS using monthly data obtained in the first wave of the panel. The model borrows strength from data observed in preceding periods and from other domains, resulting in a remarkable reduction of the standard error compared with the generalized regression estimates. This enables us to produces monthly figures instead of an average of the preceding three months with the advantage of improved timeliness and more realistic monthly figures.

Longitudinal Estimation in Dual-Frame Surveys

❖ Yan Lu, Arizona State University, 1016 S Mariana St Apt 4, Tempe, AZ 85281, *ylu11@asu.edu*; Sharon L. Lohr, Arizona State University

Key Words: multiple frame surveys, missing data, gross flows

Traditionally, large surveys use a single sampling frame from which the sample is selected. As the population and methods used to collect survey data change, single frame surveys may miss parts of the population. In order to obtain better coverage of the population of interest and cost less, a number of surveys employ dual frame surveys, in which independent samples are taken from two overlapping sampling frames. Some current surveys follow the same households at regular time intervals so that longitudinal quantities such as transitions in employment status can be studied. In this research, statistical methods for analyzing longitudinal quantities from dual frame surveys are developed. A jackknife method is applied to estimate the variance of the estimators. The methods are applied to real datasets.

Joint Modeling of Longitudinal Binary Outcome and Time to Drop-Out: An Application Using Complex Survey Data

❖ Punam Pahwa, University of Saskatchewan, 103 Hospital Drive, Saskatoon, SK S7N OW8 Canada, *pup165@mail.usask.ca*; Chandima Karunanayake, University of Saskatchewan

Key Words: Generalized estimation equations, Bootstrap, Dropout, Complex Survey, Joint Model

Statistics Canada National Population Health Survey (NPHS) data were used to trace longitudinal changes in the mental health of immigrant women. In the NPHS a member from each household was randomly chosen to participate in this longitudinal survey. We explored the moderating affects of different risk factors on immigrant women's mental distress, which was categorized as a dichotomous variable: no/low distress vs. moderate/high distress. Regression coefficients were estimated by using the weighted generalized estimating equations approach which accounts for within-subject correlation and variance estimation was conducted by taking into account the complexities of multi-stage design using bootstrap weights computed by Statistics Canada. An alternative approach based on the joint model for a longitudinal dichotomous outcome and the time to dropout is under process.

A Method for Bias-Reduction of Sample-Based MLE of the Autologistic Function

❖ Steen Magnussen, Canadian Forest Service, 506 West Burnside Rd, Pacific Forestry Center, Victoria, BC V8Z 1M5 Canada, *steen.magnussen@nrcan.gc.ca*; Robert Reeves, Queensland University of Technology

Key Words: autologistic function, cluster sampling, maximum likelihood, bias, bias correction

Sample-based MLE of the autologistic function are biased due to the lack of independence of sampling units (clusters). An earlier study (Magnussen and Reeves 2007, Journal of Applied Statistics, in press) quantified the bias and the properties of sample-based MLE estimates in an extensive simulation study. Adding a buffer of one row of ultimate sampling unit around each sampled unit (cluster) with 'missing data' and assuming that these buffered units are mutually independent a less biased MLE of the autologistic function can be obtained. The bias reduction depends on the number of ultimate units (m) in a square sampling unit. For m = 4 adding a buffer lowered bias from 8.1% to 6.1% while for m = 7 it was lowered from 5.3% to 4.3%. We recommend the buffering since it is easy to implement without adding significantly to the computational effort.

515 Applications in Public Health and Maternal and Child Health ●

Section on Statistics in Epidemiology, Section on Health Policy Statistics

Thursday, August 2, 8:30 am–10:20 am

Ranking Children Vaccination Coverage Curves with the National Immunization Survey (NIS)

❖ Zhen Zhao, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, MS E62, Atlanta, GA 30333, *zaz0@cdc.gov*; Philip J. Smith, Centers for Disease Control and Prevention

Key Words: Ranking, Curves, AUCC, Vaccination Coverage, State

We investigated ranking vaccination coverage curves in the 4:3:1:3:3 series using data from the 2000-2005 NIS. The area under coverage curve (AUCC) was chosen as the summary statistic for ranking curves. A test statistic is derived for comparing curves and is based on the linear combination of coverage estimates used to estimate the AUCC. Ranking analyses identified the top 10 and bottom 10 states/cities, among 72 states/cities sampled by the NIS, which possess consistently high or low vaccination coverage across the 6 years with national coverage curve as reference. Massachusetts and City of Detroit have the highest and lowest coverage curve, respectively. The top ten states/cities are significantly different in vaccination coverage from the bottom ten states/cities. The matrix for both comparisonwise p-value and the experimentwise adjusted p-value was obtained.

Bayesian Hierarchical Models to Evaluate Translational Research: Connecticut Collaboration for Fall Prevention

❖ Terrence Murphy, Yale University, 6 Hunting Ridge, School of Medicine, Hamden, CT 06518, *terrence.murphy@yale.edu*; Heather G. Allore, Yale University; Mary Tinetti, Yale University

Key Words: spatio-temporal analysis, Bayesian hierarchical models, translational research, study design, falls, intervention

Second stage translational studies are necessary and difficult to evaluate. A quasi-experimental design is used to compare the rate of fall-related health care utilization of two geographically disparate areas in Connecticut to evaluate an intervention designed to reduce fall-related injuries among older persons. This evaluation examines only the two baseline years prior to intervention. The experimental units are zip code tabulation areas for which data is gathered from public health sources. We estimate associations with sex, age, treatment, year of study and treatment-by-year interaction using spatio-temporal Bayesian hierarchical models. After adjusting for spatial variability, we find no association between rate of utilization and intervention and no difference in the temporal patterns of the rate of the intervention and usual care study arms in the two year baseline period.

Disparity of Mortality from Leading Causes by Socioeconomic Status and Their Geographic Distribution: U.S. Residents, 2000-2004

❖ Jay H. Kim, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *jkim@cdc.gov*

Key Words: Mortality Rates, Socioeconomic Status, Health Service Areas, Analysis of Variance, Principal Component Method

Declining mortality rates from heart diseases and cancer have still remained as leading causes of deaths. The mortality rates from other leading causes—cerebrovascular disease, chronic obstructive pulmonary disease and accidents—have either increased or remained the same. In this paper, I plan to map the distribution of mortality for five leading causes of death by health service areas that are based on the hospital usages by Medicare patients. The results are mapped using the mapping software, ArcGIS. In addition, the relationship of the changing mortality by various socioeconomic status (SES) are explored using 2000–2004 mortality data. SES was constructed by applying principal component method to socioeconomic variables available from the 2000 U.S. census. The differences in mortality rates in relation to high/low SES are evaluated by using the analysis of variance method.

Alternative Models of Calculating Risk of Heat Death Among Individuals Making Covert Border Crossings from Mexico to Arizona

❖ Mary Mays, Arizona State University, 4747 E Olney Ave, CONHI, Higley, AZ 85236, *mary.mays@asu.edu*; Samuel M. Keim, Arizona State University; Kendall Allred, Arizona State University; Jan Jirsak, Arizona State University

Key Words: estimating risk, risk communication, logistic regression, receiver operating characteristic curve, heat death, illegal immigration

Background: Heat deaths were rare in Pima County, AZ prior to 2000, but increased steadily from 10 in 1999 to 75-110 in 2002-2005. Data from 2002-2005 indicate that 90% of heat deaths occurred among individuals crossing the Mexico/U.S. border covertly. Defining the relationship between ambient temperatures and risk of heat death is difficult, because the number of covert border crossers can not be measured. Method: Risk of heat death when ambient temperatures exceed 32C was calculated using two different models for estimating the at-risk population, one based on people and one based on opportunity. Five methods were used for evaluating risk: proportions, odds ratios, logistic regression, probit analysis, and receiver operating characteristic curves. Results: People OR = 19.7; Opportunity OR = 19.3. Conclusion: A meaningful estimate of risk can be calculated using an opportunity model.

A Strategy To Evaluate Noncoverage in a Binational, Population-Based Maternal and Child Health Survey on the U.S.-Mexico Border

❖ Ruben Smith, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mailstop K21, Atlanta, GA 30341, *eyb4@cdc.gov*; Jill McDonald, Centers for Disease Control and Prevention; Christopher Johnson, Centers for Disease Control and Prevention

Key Words: data linkage, birth certificate, maternal and child health data

A design for collecting standardized reproductive health surveillance data from mothers of live-born infants was tested in two sister communities on the US-Mexico border. Mothers were sampled from six Mexican and four U.S. hospitals. To assess the potential bias of excluding mothers who delivered outside study hospitals (7% in Mexico, 2% in the US), we compared birth certificate (BC) characteristics of study hospital infants to all infants born during the study period and found no differences. Using multi-variable linkage, we matched sample data to BCs to estimate non-coverage and assess potential bias in the sampling frame. Eighty-seven percent of Mexican and 97% of U.S. women sampled were successfully matched to a BC. Three percent of both Mexican and U.S. BCs could not be linked to sample data. Results indicate minimal survey noncoverage.

Latent Variable Models for Development and Validation of an Infant Morbidity Index

❖ Xuefeng Liu, Wayne State University, 2370 Golfview Drive, Apt# 103, Troy, MI 48084, *xli@med.wayne.edu*; Jeffrey Roth,

Key Words: infant morbidity outcomes, latent variable model, modified Gauss-newton algorithm, validation, infant morbidity index

Birth Defect (BD), Abnormal Conditions (AC), Developmental Delay or Disability (DDD) and Low Birth Weight (LBW) are four major pregnancy outcomes which are associated with infant morbidity. Little attention has been paid to development of a composite index which is a summary construct of infant morbidity outcomes. In this paper, a new single variable, called infant morbidity index (IMI) which is a summary of these four infant morbidity outcomes, was developed by using latent variable models (LVM) and modeling conditional probability as a function of IMI in LVM. The validity of this index was then assessed in detail. It was shown that IMI was correlated with each of the individual outcome, with infant mortality and with a face-valid index of morbidity outcomes, and could be used in future research as a measure of infants propensity for morbidity.

Birth Weight Distribution: Thinking Outside the Curve

❖ Richard Charnigo, University of Kentucky, 851 Patterson Tower, Lexington, KY 40506-0027, *richc@ms.uky.edu*; Lorie W. Chesnut, The University of Alabama at Birmingham; Tony LoBianco, University of Kentucky

Many challenging questions about birth weight, gestational age, and mortality have arisen in recent perinatal research. To address them requires a realistic characterization of a population's birth weight distribution. We propose using a normal mixture model in which the number of components is estimated from the data and can vary across populations. In our analysis of data from several populations, typically a four-component model was preferred; the first component mean fell in the extremely low birth weight range, while the other three component means fell in the normal range. This pattern and notable exceptions to it motivate further research to describe the mixture components in terms of observable covariates and to assess how birth weight-specific mortality varies between and within components.

516 Biometrics: A Gold Standard? How Can Statistics as a Field Contribute Toward Finding an Answer? ●

Section on Statisticians in Defense and National Security, ENAR, Biometrics Section, Section on Health Policy Statistics

Thursday, August 2, 10:30 am–12:20 pm

Biometric Recognition: Applications, Performance, and Challenges

❖ Anil K. Jain, Michigan State University, Department of Computer Science and Engineering, 3115 Engineering Build., East Lansing, MI 48824, *jain@cse.msu.edu*

Key Words: security, fingerprint matching, face recognition, error rate, fusion, encryption

Biometric recognition refers to the automatic recognition of individuals based on their physiological or behavioral traits (e.g., fingerprint, face), thus allowing us to confirm or establish identity based on “who he is”, rather than by “what he possesses (ID card) or remembers” (password). Despite better security, biometric systems have to contend with problems related to non-universality of the trait, limited degrees of freedom, large intra-class variability and spoof attacks. Although the first Automatic Fingerprint Identification System (AFIS) was installed around 1965, biometric recognition remains a difficult pattern recognition problem. In this talk, we present an overview of biometrics, its advantages and limitation, state-of-the-art performance figures and highlight the role of Statistics in addressing challenges in dealing with accuracy, individuality, and security issues.

Statistical Models for Fingerprint Individuality

❖ Sarat Dass, Michigan State University, Department of Statistics and Probability, A430 Wells Hall, East Lansing, MI 48824, *sdass@stt.msu.edu*; Yongfang Zhu, Michigan State University; Anil K. Jain, Michigan State University

Key Words: Mixture models, Clustering, Biometric Authentication, Fingerprint Individuality

Forensic evidence based on fingerprints was first challenged in the 1999 case of USA vs. Byron Mitchell, and subsequently, in 20 other cases involving fingerprint evidence. The main concern is the lack of scientific validation of fingerprint evidence. The probability that two different individuals will share a common set of fingerprint features are currently unknown or unsatisfactory. This talk develops a family of finite mixture models to represent the distribution of minutiae in fingerprint images, including minutiae clustering tendencies and dependencies in different regions of the fingerprint image domain. We show that the proposed models better describe the observed variability in the minutiae compared to the uniform model reported in the literature. A mathematical model that computes the probability of a random correspondence is developed based on the mixture models.

Information Fusion in Biometrics

❖ Arun Ross, West Virginia University, PO Box 6109, Morgantown, WV 26506, *arun.ross@mail.wvu.edu*

Key Words: Biometrics, Fusion, Multibiometrics, Correlation, Imbalance, Diversity

Multibiometrics refers to the process of consolidating the evidence presented by multiple biometric sources in order to enhance the recognition accuracy of a biometric system. Thus, multibiometric systems combine the information presented by multiple biometric sensors (e.g., thermal/IR cameras), algorithms (e.g., minutiae-/ridge-based fingerprint matchers), samples (e.g., frontal-/side-profile fae images), units (e.g., left/right irises) or traits (e.g., face/fingerprint). This work examines the various levels of fusion as well as the different fusion techniques that have been described in the literature. The impact of correlated biometric sources on the performance of different fusion rules at the score level is analyzed. Experimental results help in understanding the role of correlation, diversity and imbalance between biometric sources on the performance of different fusion rules.

Statistical Issues in Biometric Authentication

❖ Stephen Fienberg, Carnegie Mellon University, Department of Statistics, School of Computer Science, Pittsburgh, PA 15213, *fienberg@stat.cmu.edu*

Key Words: biometrics, statistical analysis, ROC curve, empirical methods, error rates

A wide variety of biometric authentication systems exist today, based on face, fingerprints and even iris recognition. Following the attacks of 9/11, “biometrics” migrated in use in airports and other public places, and they are about to become an integral part of the US immigration and passport systems. Many biometric systems report surprisingly high error rates and there is an acute lack of rigorous statistical methods for analyzing these. We discuss examples of biometric authentication tools and systems and ways to assess their reliability and accuracy.

517 New Development in the Analysis of Medical Data with Complex Structures ●

Biometrics Section, SSC, ENAR

Thursday, August 2, 10:30 am–12:20 pm

Multivariate Interval-Censored Survival Data: Parametric, Semiparametric, and Nonparametric Models

❖ Philip Hougaard, Lundbeck, Biostatistics, Ottiliavej 9, Valby, International DK2500 Denmark, *phou@lundbeck.com*

Key Words: multivariate, survival, interval censoring, frailty, dependence, semi-parametric

Interval censoring means that an event time is in an interval (L,R], with L the last examination time before the event, and R the first after. In the univariate case, parametric models are easily fitted, and for non-parametric models, the mass is placed on some intervals, derived from the L and R points. Asymptotic results are simple for the former and complicated for the latter. Parametric models extend easily to multivariate data, like eruption times for teeth, examined at visits to the dentist. However, nonparametric models are intrinsically more complicated. It is difficult to derive the intervals with positive mass and estimated interval probabilities may not be unique. A semiparametric model makes a compromise, with a parametric model, like a frailty model, for the dependence and a nonparametric model for the marginal. I will compare and discuss these three models.

Robust Estimation of State Occupancy Probabilities with Interval-Censored Multistate Data

❖ Richard J. Cook, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1 Canada, *rjcook@uwaterloo.ca*; David Tolusso, University of Waterloo

Key Words: interval-censoring, multi-state data, robust estimation, state occupancy

Interest frequently lies in estimation of state occupancy probabilities with multistate data. We consider robust methods of estimating occupancy probability functions with progressive multistate data where the transition times are interval censored. An emphasis is on weakly parametric models based on Markov assumptions and based on differences in marginal survivor functions. Robustness to semi-Markov models and dependent assessment times are studied. The methods are illustrated by application to data from a cancer clinical trial.

A Class of Linear Transformation Models Under Interval Censoring

❖ Zhigang Zhang, Oklahoma State University, 301C MSCS, Dept. of Statistics, Stillwater, OK 74078, *zhigang.zhang@okstate.edu*

Key Words: Linear transformation models, interval censoring, regression analysis

Regression analysis of interval-censored failure time data has recently attracted a great deal of attention and for the problem, several methods have been proposed. However, most of these methods employ specific models such as the proportional hazards model. The linear transformation models, a class of generalized proportional hazards models, provide great flexibility in fitting survival data. Although the application of these models on right-censored data has been discussed, the only existing inference procedure for the linear transformation models when interval-censored data are observed is limited to multi-sample situations and lacks rigorous technical justification. In this study we present approaches that allow the covariate to be arbitrary and provide a theoretical framework.

Inference After Model Selection Using Restricted Permutation Methods

❖ Rui Wang, Harvard School of Public Health, 655 Huntington Ave.,, BLDG II, Room 428, Boston, MA 02115, *rwang@hsph.harvard.edu*; Stephen Lagakos, Harvard School of Public Health

Key Words: automated variable selector, covariates, regression, training and validation data sets

Many statistical problems first apply a model-selection algorithm to identify a subset of candidate covariates, and then make inferences about the association between the selected covariates and the response. It is well-known that ignoring the fact that the covariates were selected on the basis of their apparent association with the response can lead to biased inference. We develop inference methods based on partitioning the response-covariate data matrix in specific ways for purposes of generating permutations, depending on the (arbitrary) model selection procedure and correlation structure of the response-covariate data matrix, and then basing inferences on a restricted set of permutations. We illustrated how the proposed methods correct biases from model selection, their efficiency relative to use of an independent validation set, and their robustness to violations of assumptions.

518 Wald Lecture III

IMS, General Methodology, Section on Bayesian Statistical Science

Thursday, August 2, 10:30 am–12:20 pm

Working with Inexact Models: The World of Computer Modeling

❖ James Berger, Duke University, Inst of Statistics and Decision Science, Durham, NC 27708-0251, *berger@stat.duke.edu*

Key Words: math models, prediction, confounding, modular Bayes, spatial, functional data

A major activity in science and engineering is the development of simulation- or math-based computer models of processes. Such models are virtually always incomplete representations of reality and hence will be rejected by formal tests. The models will be used, however, so the statistical challenge is to understand how to do so effectively. Statistical issues that arise include determination of model bias, use of bias-adjusted predictions and accounting for uncertainty in model parameters and inputs. The methodology used is a mix of Bayesian spatial, hierarchical and nonparametric techniques; such techniques seem necessary, as will be shown with a simple example. Difficulties encountered include needing to deal with functional data, the presence of severe confounding and the surprising phenomenon that full Bayesian analysis is often inferior to a modular, or partial, Bayesian analysis.

519 Recent Advances in Group Randomization Trials ● ✪

WNAR, ENAR, Section on Health Policy Statistics,
Biometrics Section

Thursday, August 2, 10:30 am–12:20 pm

Modeling Heterogeneity in Intraclass Correlation

❖ Catherine M. Crespi, University of California, Los Angeles, Department of Biostatistics, CHS 51-254, Los Angeles, CA 90095-1772, *ccrespi@ucla.edu*; Weng-Kee Wong, University of California, Los Angeles

Key Words: Generalized estimating equations, Correlated data, Group randomized trials, Intraclass correlation

In cluster-randomized trials, it is commonly assumed that the magnitude of the correlation among subjects is invariant across clusters. However, the correlation may in fact vary among clusters due to the delivery of the intervention or cluster-level covariates. Failure to adequately model heterogeneity of the correlation may result in loss of efficiency. We present a method for modeling heterogeneous correlation structure using generalized estimating equations and demonstrate the use of goodness of fit criteria for choosing the best correlation structure. Application to a cluster-randomized trial of an educational intervention to promote cancer screening illustrates the methods.

Breaking the Matches in Community Intervention Trials

❖ Allan Donner, University of Western Ontario, 1151 Richmond St., London, ON N6A 5C1 Canada, *allan.donner@schulich.uwo.ca*

Key Words: community intervention trials, pair-matching, regression analysis, clustering

Pair-matched cluster randomization trials are frequently adopted as the design of choice for evaluating an intervention offered at the community level. However previous research has demonstrated that a strategy of breaking the matches and performing an unmatched analysis may be more efficient than performing a matched analysis on the resulting data, particularly when the total number of communities is small and the matching is judged as relatively ineffective. The research concerning this question has naturally focused on testing the effect of intervention. However an important secondary objective of many community intervention trials is to investigate the effect of individual-level risk factors on one or more outcome variables. Focusing on the case of a continuous outcome variable, we investigate the impact of breaking the matches from the perspective of this objective.

Contamination and Group Randomized Trials

❖ Ziding Feng, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M2-B500, Seattle, WA 98109, *zfeng@fhcrc.org*

Key Words: Contamination, Group Randomized Trials

Investigators often automatically use group randomized trials when faced with the possibility of contamination across subjects within a group. However, the bias due to contamination always makes the observed intervention effect smaller, leading to conservative rather than spurious findings. Therefore the efficiency is the central issue. It has been suggested (Torgerson, 2001) that individual randomized trials may nevertheless be more efficient despite the contamination. We present models for contamination that allow quantification of the dilution of effect size versus the lower cost of individual randomized trials as a function of GRT design parameters, such as the ICC. This will allow calculation of the degree of contamination necessary to make a GRT more efficient and aid researchers in avoiding GRTs when they are unnecessary and not cost-effective.

Design and Analysis of Pseudo-Cluster Randomized Trials

❖ Mirjam Moerbeek, Utrecht University, PO Box 80140, Utrecht, International 3508 TC Netherlands, *m.moerbeek@fss.uu.nl*

Key Words: cluster randomization, treatment group contamination, selection bias, design effect

The main methodological reason for cluster randomization is to avoid treatment group contamination, which may occur when persons are randomized to treatments. Disadvantages of cluster randomization are reduced efficiency, selection bias, and a slow recruitment in the clusters that are randomized to the “less interesting” treatment. Pseudo cluster randomization has been proposed as a compromise between person and cluster randomization. This presentation provides sample size formulae and the design effect. Furthermore, it compares different estimation methods on basis of a simulation study, namely GEE, mixed models, and an estimator originally proposed for pseudo cluster randomized trials. Mixed models without the use of weights is recommended as the best compromise between feasibility, power, and robustness.

Group Randomized Trials in Ethnic Populations

❖ Susan L. Stewart, University of California, San Francisco, 185 Berry St., Lobby 2, Suite 6600, Box 0981, San Francisco, CA 94143-0981, *sstewart@cc.ucsf.edu*

Key Words: study design, controlled trials, community intervention, clustering,, minorities

Many interrelated considerations go into the planning of a community intervention trial, including the nature of the desired outcome, the content and method of delivery of the intervention, and the characteristics of the target population. Geographical and social clustering in ethnic populations can present unique challenges and opportunities for study designers. For instance, media interventions are designed to impact an entire community, and the feasibility of randomizing communities may be adversely affected by geographical clustering. On the other hand, small group interventions can take advantage of pre-existing social clusters that facilitate randomization. In virtually all study designs, the impact of social networks on inter-cluster communication, and therefore possible contamination, must be considered. Examples of study designs that address these issues are presented.

520 Topics on Sample Survey Design and Inference in Honor of J. N. K. Rao’s 70th Birthday and 50 Years of Contributions ●

Section on Survey Research Methods
Thursday, August 2, 10:30 am–12:20 pm

Recent Developments in Multiple-Frame Surveys

❖ Sharon L. Lohr, Arizona State University, Department of Mathematics and Statistics, Tempe, AZ 85287-1804, *sharon.lohr@asu.edu*

Key Words: complex survey, dual frame survey, sampling for rare events, misclassification

With the increasing demographic and technological diversity of the U.S. population, it is becoming more difficult for a single sample selected from a single sampling frame to adequately represent the population. Multiple-frame surveys are increasingly used in situations where several sampling frames may provide better coverage or cost-efficiency for estimating population quantities of interest. Examples include combining a list frame of farms with an area frame or using two frames to sample landline telephone households and cellular telephone households. We review the history of multiple frame surveys including J.N.K. Rao’s many contributions to the subject. We then discuss some recent work on internally consistent and efficient estimators for three or more frames, effects of frame misclassification, and resampling methods for variance estimation in multiple frame surveys.

Rao-Scott Corrections and Beyond

❖ Alastair J. Scott, University of Auckland, Department of Statistics, 38 Princes Street, Auckland, 1010 New Zealand, *scott@stat.auckland.ac.nz*

Key Words: Survey Data, Chi-squared Tests

It is now 25 years since Rao-Scott corrections to chi-squared tests were introduced for two-way tables (Rao & Scott 1981) and more than 20 years since the extension to multi-way tables (Rao & Scott 1984) was published. Until five years ago, both papers were attracting one or two citations per year but this has increased rapidly to a rate of 20 or so per year over the last five years. The primary reason for this is the inclusion of the corrections as standard features in most big computer packages for the analysis of survey data. We look at computational aspects of the methods and at some improvements and extensions. References. Rao, J.N.K. & Scott, A.J. The analysis of categorical data from complex sample surveys. JASA, 76, 1981, 221-230. Rao, J.N.K. & Scott, A.J. On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. AnnStat,12, 1984, 46-60.

Small-Area Estimation: Theory and Practice

❖ Mike A. Hidiroglou, Statistics Canada, Section D, 16 th Floor R.H. Coats Building, Tunney’s Pasture, Ottawa, ON K1A 0T6 Canada, *hidirog@yahoo.ca; Mike.Hidiroglou@statcan.ca*

Key Words: Small, Area, Official, Statistics

Small-area estimation was first studied at Statistics Canada in the seventies. Small-area estimates have been produced using administrative files or surveys enhanced with administrative auxiliary data since the early eighties. In this paper we provide a summary of existing procedures for producing official small-area estimates at Statistics Canada, as well as a summary of the ongoing research. We discuss the estimation of under-coverage in the census when small-area estimation techniques are used. We also highlight problems for producing small-area estimates for business surveys.

521 Some Practical Considerations for Statisticians Designing Clinical Trials ●

Statisticians in the Pharmaceutical Industry, ENAR, Biometrics Section, Biopharmaceutical Section

Thursday, August 2, 10:30 am–12:20 pm

Statistical Considerations in Planning Drug Clinical Trials

❖H.M. James Hung, Food and Drug Administration, 10903 New Hampshire Ave Bldg 22 Rm 4238, HFD-710 Mail Stop #4105, Silver Spring, MD 20993-0002, *hsienming.hung@fda.hhs.gov*

Key Words: sample size, study-wise type I error, effect size, reproducibility

The usual regulatory evidential requirement for efficacy of a test drug is that at least two clinical trials provide sufficient statistical evidence of the therapeutic effect. Statistical significance for the estimated drug effect is dependent on the planned sample size for detecting the postulated effect size with a certain level of statistical power. Reproducibility of a statistically significant finding needs to be factored into the planning of the two registration trials. The postulated effect size is related to the effect size estimated from other trials. Moreover, per the current standard, a typical registration trial requires that the overall study-wise type I error rate be no larger than 5%. In this paper, we will discuss the relationships between these factors and share our insights into the implications of these considerations on the trial planning.

Performance of Multiple Testing Procedures To Compare Three Doses of a Test Drug and Placebo

❖Lan Kong, University of Pittsburgh, 313 Parran Hall, 130 Desoto Street, Pittsburgh, PA 15261, *lkong@pitt.edu*; Gary Koch, The University of North Carolina at Chapel Hill; Thomas Liu, Amgen Inc.; Huei Wang, Amgen Inc.

Key Words: closed testing procedure, familywise Type I error, multiple comparison, power

A study design with two or more doses of a test drug and placebo is frequently used in clinical drug development. Multiplicity issues arise when there are multiple comparisons between doses of test drug and placebo, and also when there are comparisons of doses with one another. An appropriate analysis strategy needs to be specified in advance to avoid spurious results through insufficient control of Type I error, as well as to avoid the loss of power due to excessively conservative adjustments for multiplicity. For evaluation of alternative strategies with possibly complex management of multiplicity, we compare the performance of several testing procedures through the simulated data that represent various patterns of treatment differences. The purpose is to identify which methods perform better or more robustly than the others and under what conditions.

Choice of the Primary Analysis in Longitudinal Clinical Trials

❖Craig Mallinckrodt, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IA 46285, *cmallinc@lilly.com*

Key Words: missing data, longitudinal data, clinical trials

The last observation carried forward (LOCF) approach is often used to handle missing data. Advances in methodology and its implementation have made likelihood- and Bayesian-based approaches easy to implement and they are commonly used. Research has unequivocally demonstrated that these approaches are more robust to the biases from missing data and provide better control of Type I and Type II errors than LOCF. The supposed conservative behavior of LOCF is never guaranteed and is often unlikely. Furthermore, using LOCF as a composite measure of safety, tolerability and efficacy can lead to erroneous conclusions regarding the effectiveness of a drug. Although practice is shifting away from LOCF, regulatory acceptance of alternative approaches is slow. Greater acceptance of MAR-based approaches is needed to improve the efficiency and effectiveness of drug development.

Bioequivalence and the Pharmaceutical Industry

❖Scott Patterson, GlaxoSmithKline, 2301 Renaissance Boulevard, King of Prussia, PA 19406, *scott.d.patterson@gsk.com*; Byron Jones, Pfizer Inc.

Key Words: bioequivalence, REML, method-of-moments

Average bioequivalence (ABE) studies have long been the regulatory standard for demonstrating that two formulations of drug product will provide the same benefit and safety profile when used in the market. Population (PBE) and Individual (IBE) BE were the subject of debate when proposed in FDA guidance in 1997. This called for the use of replicate design, crossover studies for selected drug products and REML models for parameter estimation. Market access would be granted if the products demonstrated ABE, and sponsors had the option of using PBE or IBE if the use of these criteria could be justified. SAS-based approaches to the modeling of pharmacokinetic data from such studies will be summarized based on work published by the authors in 2002 and 2004 in Pharmaceutical Statistics, updated with recent findings.

522 National Science Foundation–Funded Projects in Undergraduate Education

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Thursday, August 2, 10:30 am–12:20 pm

The SUM Project: A Statistics Undergraduate Mentoring (SUM) Program

❖Julie Legler, St. Olaf College, Old Music Hall 101, 1520 St Olaf Ave, Northfield, MN 55057, *legler@stolaf.edu*

Key Words: Education, Undergraduate, Collaborative research, Curriculum

The Statistics Undergraduate Mentoring (SUM) Project will demonstrate how four-year institutions can successfully recruit and mentor undergraduate and post-doctoral associates in statistics. The project goals are to increase the number of undergraduates from four-year colleges who attend graduate school in statistics and to increase the number of PhDs in statistics who pursue careers at four-year colleges. A Center for Interdisciplinary Research (CIR) was established where undergraduates, post-

doctoral associates, and faculty mentors collaborate on statistical aspects of professional and undergraduate research with faculty and students from other departments on campus. Post-doctoral associates assist in running the CIR, teach a variety of undergraduate courses, work with students in inter-disciplinary research projects, and continue to develop their own research program.

Content-Rich Statistical Projects for Undergraduate Economics Students

❖ Linda Collins, The University of Chicago, Department of Statistics, 5734 S. University Avenue Eckhart 107, Chicago, IL 60637, *collins@galton.uchicago.edu*

Key Words: Education, Undergraduate, Curriculum, labs, economics, projects

We present statistical projects intended for undergraduate economics students with a statistics prerequisite. Projects model the interplay of statistics and economics needed for successful economic modeling, keep statistical thinking at the forefront, but also expect students to include economic theory in the model-building process. Projects generally include background reading from the economics literature that includes discussion of statistical approaches. Students should have some exposure to economic theories and some mathematical maturity in order to investigate mathematical underpinnings of the statistical methods, which may include variable selection in multiple regression, modeling time- or spatially-dependent data, Bayesian methods and general linear models. If time, will present support materials for students applying mathematics knowledge to new problems for the first time.

Computing in Statistics: Model Courses and Curricula

❖ Deborah Nolan, University of California, Berkeley, Department of Statistics, 367 Evans Hall #3860, Berkeley, CA 94720-3860, *nolan@stat.berkeley.edu*; Mark Hansen, University of California, Los Angeles; Duncan Temple Lang, University of California, Davis

Key Words: Education, Undergraduate, Curriculum, Computing

The ability of students to compute has a real impact on their ability to understand and explore statistical concepts and to contribute significantly to the study of important scientific problems. This project aims to develop course materials and sample curricula that expand the role of computing in undergraduate statistics programs. The first phase of the project brings together educators to discuss ideas for how to introduce new courses in computing and integrate computing into traditional statistics courses. The areas of focus are emerging data technologies, visualization, programming languages and environments, and statistical algorithms. The project also provides forums to discuss and disseminate advances in technology and their impact on statistics and workshops to both teach faculty about these topics and demonstrate how to teach these topics in undergraduate programs.

Interdisciplinary Labs Created for a First or Second Course in Statistics

❖ Shonda Kuiper, Grinnell College, Department of Mathematics and Statistics, Science 2407, Grinnell, IA 50112-1690, *KUIPERS@grinnell.edu*

Key Words: Education, Undergraduate, Curriculum, Labs, On-line experiments

This paper presents multi-day lab modules that emphasize the process of science and data analysis relevant for science and social science students. These investigative labs address relatively advanced statistical techniques and encourage students to collect data, determine an appropriate technique for analysis, integrate technology, perform the analysis, make inferences and interpret the results. In addition, they encourage students to experience the role of a research scientist in searching the literature, preparing a proposal for analysis, planning and carrying out experiments, and presenting the results. These labs can be combined to form a second statistics course, individually incorporated into an introductory statistics course, or used in courses in other disciplines.

523 Dimension Reduction and Information Visualization ● ✪

IMS, Section on Nonparametric Statistics, Section on Statistical Graphics

Thursday, August 2, 10:30 am–12:20 pm

An RKHS Formulation of the Inverse Regression Dimension Problem

❖ Tailen Hsing, The Ohio State University, Department of Statistics, Columbus, OH 43210, *hsing@stat.ohio-state.edu*; Haobo Ren, Alcatel-Lucent

Key Words: functional data, nonparametric regression, inverse regression, reproducing kernel Hilbert space

Suppose that Y is a scalar and X is a second-order stochastic process, where Y and X are conditionally independent given the random variables $\xi_1,\ldots,\xi_p$ which belong to the closed span L_X^2 of X. This paper investigates a unified framework for the inverse regression dimension reduction problem. It is found that the identification of L_X^2 with the reproducing kernel Hilbert space of X provides a platform for a seamless extension from the finite to infinite-dimensional settings. It also facilitates a convenient computational algorithm that can be applied to a variety of models.

Iterative Robust Sliced Inverse Regression for Gene Clustering in Microarray Gene Expression Data Analysis

❖ Han-Ming Wu, Academia Sinica, Institute of Statistical Science, 128 Academia Rd. Sec.2, Nankang, Taipei, 11529 Taiwan, *hmwu@stat.sinica.edu.tw*

Key Words: Dimension reduction, K-means clustering, unsupervised learning

Clustering techniques such as hierarchical clustering, K-means are useful in revealing biologically relevant groups of genes in microarray data analysis. However, due to the noisy nature of microarray data, these methods are sensitive to outliers. In this study, an iterative robust version of sliced inverse regression (SIR) (Li, 1991) is proposed. Firstly, the partitioning around medoids (PAM) is used to generate the initial cluster labels. Then the robust version of SIR is combined with the nearest medoids classifier (NMC) to iteratively update the cluster labels for genes. Simulation results and real microarray data analyses show that the proposed method is capable of identifying more coherent expression patterns and providing more stable clusters than the K-means method.

Matrix Visualization for High-Dimensional Categorical Data Structure with a Cartography Link

❖ Chun-houh Chen, Academia Sinica, Institute of Statistical Science, 128 Academia Road, Section 2, Taipei, International 115 Taiwan, *cchen@stat.sinica.edu.tw*; Sheng Li Tzeng, Academia Sinica; Chiun-How Kao, Academia Sinica

Key Words: categorical data, color coding, homogeneity analysis, geographic visualization, proximity measure

Matrix visualization (MV) is more efficient than conventional graphical tools such as scatterplot, boxplot, and parallel-coordinate-plot in extracting information structure embedded in moderate (hundreds) to high dimensional continuous data. For non-continuous data, conventional tools can not provide much visual information while MV gives us information about individual profiles with interaction of subject-clusters on variable-groups. When an cartography link is attached to each sample of a high-dimensional categorical data matrix, it is desired to use a geographical map to illustrate the pattern of subject (region)-clusters with variable-groups embedded in the high-dimensional space. This study presents an interactive cartography system with systematic color coding by integrating the homogeneity analysis into matrix visualization. For more information see: http://gap.stat.sinica.edu.tw.

524 Recent Advances in Nonstationary Time Series

Business and Economics Statistics Section, IMS

Thursday, August 2, 10:30 am–12:20 pm

Local Adaptive Estimation for Nonstationary Time Series

❖Sebastien Van Bellegem, Universite catholique de Louvain, Institut de Statistique, Voie du Roman Pays, 20, Louvain-la-Neuve, B-1348 Belgium, *vanbellegem@stat.ucl.ac.be*

Key Words: Adaptive estimation, Locally stationary processes, Structural Break, Unit Root, Time-varying parameters

We present a new method to fit locally a stationary model to a nonstationary process. The method is based on a sequence of local tests of homogeneity, to detect segments of data that are nearly stationary. We establish some results concerning the optimality of the procedure, under mild assumptions. We also study the performance of the procedure in particular situations, such as time series with structural breaks or with a smooth evolution in their autocovariance function over time. An application of the procedure to econometric time series illustrates the reasonable performance of the procedure in practice.

Discrimination of Nonstationary Signals Using Localized Higher-Order Spectra

❖Hernando C. Ombao, University of Illinois, 725 S. Wright St., Department of Statistics, Champaign, IL 61820, *ombao@uiuc.edu*

Key Words: Non-linear, Non-stationary, Localized transforms, Discrimination, Classification, Brain Signals

We consider a data set that consists of MEG (magnetoencephalogram) signals recorded from healthy controls and schizophrenic patients. Our neuroscientist collaborators are interested in identifying features that can separate the two groups with the hope that these physiological measures may be used in conjunction with behavioral measures for patient diagnosis. In this talk, we will develop an automatic procedure for time-frequency spectral feature selection via localized transforms. Moreover, given the high degree of complexity of brain signals, we will consider the time-evolutionary spectrum of non-linear transforms as potential features for classification and discrimination.

Time-Dependent Spectral Analysis with Applications to Epileptic EEG Data

❖Wensheng Guo, University of Pennsylvania, 423 Guardian Drive, Blockley Hall 613, Philadelphia, PA 19104-6021, *wguo@cceb.upenn.edu*

Key Words: Locally Stationary Process, Smoothing Splines, Time-frequency Analysis, Multivariate time series, Functional data analysis, Principal component analysis

Guo et al (2003, JASA) proposed a locally stationary time series model whose time-varying transfer function is smooth in both time and frequency domain. This essentially turns the time-varying spectral estimation into a two-dimensional surface estimation problem and allows joint smoothing of the time and frequency domains. Using the smooth time-varying spectrum as the building block, the method can be extended to the multivariate setting, and to the functional data analysis setting with a family of locally stationary time series indexed by a set of covariates. In this talk, we will present the methods and their applications to epileptic EEG data.

525 Risk Estimation Challenges ● ✪

Section on Risk Analysis

Thursday, August 2, 10:30 am–12:20 pm

Risk and Regulatory Policy Assessment Based on Complex Data Sources

❖Michael E. Tarter, University of California, Berkeley, 2717 Benvenue Ave., Berkeley, CA 94705, *Tarter@berkeley.edu*; Joseph P. Brown, State of California, Air Toxicology, and Risk Assessment Unit; Andrew G. Salmon, State of California, Air Toxicology and Risk Assessment Unit

Key Words: Animal experiments, Best Linear Estimation, Order statistics, Small sized samples, Subsurvival, Toxicity

It is convenient to estimate a 95% tolerance bound for the time T beyond which 90% or more subjects had not, as yet, died with targeted tumor present. The quantity 95% applies to what is often called uncertainty, and the quantity 90% is associated with the curve property, variability. Two factors influence uncertainty, experiment quality and data quantity. A bound's low value makes a putative toxic substance appear harmful. Hence experiments conducted with an insufficient number of animals, say n, are penalized for this false economy by a low-valued bound which, in a well thought out experiment would have, as n increased, approached T. That T, not the median or some other measure of central tendency, is targeted, takes experimental subject variability into account. The less the variability, the closer T will be to the population median or to some other measure of central tendency.

Estimating Cancer Risks from Exposure to Low Levels of Ionizing Radiation

❖Ethel Gilbert, National Cancer Institute, 6120 Executive Blvd., Room 7050, Rockville, MD 20852, *gilberte@mail.nih.gov*

Key Words: radiation, dose-response, excess relative risk, excess absolute risk

A large number of epidemiologic studies of persons exposed to radiation for environmental, medical, and occupational reasons have been conducted. In many of these studies, including the important study of Japanese A-bomb survivors, well characterized dose estimates are available for individual subjects. This wealth of data has led to sophisticated statistical models for the excess relative and absolute risk. These models serve as the basis for estimating lifetime risks from low level exposures that occur in environmental, diagnostic medical, and occupational settings of interest Issues to be discussed include combining data from several relevant studies, extrapolating from high to low doses and dose rates, modification of risk by factors such as age and gender, and transporting risks developed from specific study populations to other populations with different baseline rates.

Risk and the Early Detection of Disease

Marvin Zelen, Harvard School of Public Health; ❖Sandra Lee, Dana-Farber Cancer Institute, Dept of Biostat and Comp Bio, 44 Binney St., Boston, MA 02115, *sjlee@jimmy.harvard.edu*

Key Words: Early Detection, Stochastic Model, Exam Schedules, Mammography

Currently, there are many recommendations for individuals to seek special exams to detect chronic diseases even though an individual has no signs or symptoms. The recommendations for screening exams consist of the age to begin the exams and the intervals between exams. Usually intervals between exams are fixed; e.g. every year, every two years, etc. However in nearly all adult cancers, the disease risk increases with age. This talk will discuss a method for scheduling screening exams according to risk status, which we call the "threshold" method. Illustrations will be made of the applications of the threshold method to breast and colorectal cancers compared to current screening guidelines for these diseases with regard to mortality reduction.

526 Recent Advances in Statistical Learning and Data Mining

ASA Interest Group on Statistical Learning and Data Mining, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences

Thursday, August 2, 10:30 am–12:20 pm

Probability Estimation for Large Margin Classifiers

❖Junhui Wang, Columbia University, Department of Statistics, New York, NY 10027, *jwang@stat.columbia.edu*; Xiaotong Shen, The University of Minnesota; Yufeng Liu, The University of North Carolina at Chapel Hill

Key Words: Function estimation, High dimension and low sample size, Interval estimation, Tuning, Weighting

Large margin classifiers have proven to be effective in delivering high predictive accuracy, particularly those focusing on the decision boundaries and bypassing the requirement of estimating the class probability given input for discrimination. As a result, these classifiers may not directly yield an estimated class probability, which is of interest itself. In this talk, I will present a novel method to estimate the class probability through sequential weighted classifications, by utilizing features of interval estimation of large margin classifiers. In particular, I will discuss four aspects: (1) the idea and methodology development; (2) tuning parameter selection; (3) regularization solution path; (4) a statistical learning theory. Numerical examples will be provided to demonstrate the advantage of our proposed methodology.

Screening for Monotone Association

❖Joseph S. Verducci, The Ohio State University, Dept. of Statistics, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *verducci.1@osu.edu*

Key Words: Kendall's tau, Ulam metric, non-parametric, discovery, false negatives

General procedures search through many pairs of variables that have been measured on a common sample, identifying those that are likely to have a structural relationship on some subset of the sample. An unusual aspect considered here is that a subsample is being screened along with the variable-pairs. This is motivated by as example where different subsets of cancer cell-lines may show different types of association between the biological activity and microRNA expression. Several statistical approaches include counting concordances, transforming to a location problem, gradient searching and expanding longest monotone sequence algorithms. All of these behave differently depending on the nature of the association, but for all methods the association must be quite strong to have much chance at correctly identifying the subsample of interest.

Lasso Regression: Some Recent Developments

❖David Madigan, Rutgers University, 77 Hamilton Street, New Brunswick, NJ 08901, *dmadigan@rutgers.edu*

Key Words: lasso, massive data

L1-penalized regression has attracted considerable attention in recent years. The approach provides simultaneous variable selection and shrinkage and scales to ultra high-dimensional problems. Recent developments include the group lasso and the fused lasso. This talk will review these developments. Motivated by applications in vaccine studies and early-stage drug safety modeling I will describe some recent progress combining hierarchical partition models with fused and grouped lasso within a Bayesian framework. (joint work with Suhrid Balakrishnan)

527 Recent Outreach from the American Statistical Association to the General Public

The American Statistician, Section on Statistical Education

Thursday, August 2, 10:30 am–12:20 pm

Garnering Student Interest: Finding Homer Simpson in Homework Assignments

❖Leonard A. Stefanski, North Carolina State University, Dept. of Statistics, Box 8203, Raleigh, NC 27695-8203, *stefansk@stat.ncsu.edu*

Do you want to entice your linear regression students to examine residual plots? How about added-variable plots? Do you want them to approach variable selection exercises with enthusiasm? Then instill a Pavlovian response by rewarding their efforts with a surprise at the end of their analyses. This talk will describe an algorithm for generating multiple linear regression datasets having the property that a user-specified image or message appears in the scatter plot of residuals versus fitted values (or in an added-variable plot) from the correct model fit to the data. Example datasets will be presented and their use in regression analysis and variable selection exercises illustrated. Time permitting, we'll tour a web page containing several downloadable datasets, and I'll share some experiences related to the ASA's efforts promoting this research.

Attracting Sports Enthusiasts: Is an .833 Hitter Better Than a .338 Hitter?

❖Jesse Frey, Villanova University, Department of Mathematical Sciences, 800 Lancaster Ave, Villanova, PA 19085, *jesse.frey@villanova.edu*

Key Words: Baseball; Batting average; Bayesian methods

This talk considers the problem of using batting average alone to estimate a baseball player's chance of getting a hit. This problem differs from typical proportion estimation problems because neither the number of successes nor the number of trials is known. It is also a censored data problem because batting average is reported to only three decimal places. We solve the problem in the context of present-day major league baseball by first developing a model for the joint distribution of hits, at bats, and chance of getting a hit. We then treat that model as a prior distribution and update the prior in Bayesian fashion.

Recent Public Relations Outreach of the ASA: Actions and Results

❖ Peter Westfall, Texas Tech University, Area of ISQS, Lubbock, TX 79409, *peter.westfall@ttu.edu*; Rosanne Desmone, American Statistical Association

Key Words: ASA, Promotion, The American Statistician

Recently, ASA hired a public relations specialist, Rosanne Desmone, to assist the association and our profession in achieving a more outward focus. The two recently published papers that are presented in this session were promoted vigorously through PRWeb, an online news distribution service, and eMediaWire, its newswire. In this talk, we discuss the process of going from abstract statistical content to general public interest statements; results from the promotions are also given.

528 Statistical Issues in Diagnostic and Prognostic Assays ●

Section on Statistics in Epidemiology, WNAR

Thursday, August 2, 10:30 am–12:20 pm

Time-Dependent Receiver Operating Characteristic Curves for Early Markers of Event Time Outcomes

❖ Yuying Jin, University of Washington, Department of Biostatistics, F-600 Health Sciences Building, 1705 NE Pacific Street, Seattle, WA 98195, *jinyu@u.washington.edu*; Margaret Pepe, Fred Hutchinson Cancer Research Center; Yingye Zheng, Fred Hutchinson Cancer Research Center

Key Words: time-dependent ROC, bias, efficiency, sensitivity, specificity

The received operating characteristic (ROC) curve is a popular diagnostic tool assessing the diagnostic accuracy of a putative biomarker. When a biomarker is used for prognosis, often the marker is not measured concurrently with disease outcome. Further the disease onset time can be censored. Therefore the classic concept of ROC curves needs to be extended to integrate both the time-varying nature of the marker and the clinical onset time of the disease. NaÔve estimates of predictive accuracy ignore the censoring can be biased. Incorporating the time dimension in ROC curve analysis has recently been an area of active research. We contrast the forms and attributes of various time-dependent ROC estimators discussed in the literature. The bias and efficiency of these estimators are investigated in simulation studies. Recommendations for their uses in practical situations are given.

Time Dependence of a Gene-Expression Prognosticator from Modeling Breast Cancer Survival Data

❖ Kit Fun Lau, Celera, 5319 Matthew Terrace, Fremont, CA 94502, *kit.lau@celera.com*; John Sninsky, Celera; Trevor Hastie, Stanford University

Key Words: time dependent prognostic, breast cancer, gene expression signature, time dependent AUC, metastasis score

An RT-PCR based gene-expression prognostic signature has been derived to predict distant metastasis for breast cancer. Supervised principal components and LASSO methods were used to carry out the gene selection and form a metastasis score (MS) which is a multivariate index predictor. The signature has subsequently been validated in two independent datasets. Moreover, the prognostic values of the molecular signature were shown to be most accurate for the first five years but to decrease over time. The time dependence was manifested by 1) a decrease in hazard ratios between MS high-risk and low-risk groups over time, 2) trend toward statistical significance for the time-dependence of the regression coefficients of MS in a Cox model and 3) a decrease in time-dependent AUC. Temporal evolution of the molecular prognosticator will shed light on its usefulness for different time horizons.

On Design Issues of Clinical Studies for in vitro Diagnostics

❖ Christoph Berding, Roche Diagnostics GmbH, Nonnenwald 2, Penzberg, 82377 Germany, *christoph.berding@roche.com*

Key Words: diagnostic, design, study

We review selected medical situations, where diagnostic tools support medical decision making. IVD manufacturers screen these medical areas for useful new proteomic or genomic markers or marker panels. Accordingly a new marker project starts with the formulation of a medical need and the intended use of the new diagnostic tool. Additionally clinical studies to demonstrate the performance and clinical utility of the new markers/marker panels have to be defined. Depending on project advancement these studies are either explanatory or confirmatory in nature. In early phases diagnostic studies are often retrospective case-control, which requires a careful investigation into all kinds of bias. By way of example we discuss a number of diagnostic settings and related study designs. Rather than presenting ready-to-use design solutions we would like stimulate the discussion on alternatives.

Assessment of Impact of Protocol Changes on Classification Accuracy

❖ Wei-min Liu, Roche Molecular Systems, Inc., 4300 Hacienda Drive, Pleasanton, CA 94588, *wei-min.liu@roche.com*; Wei Wen, Roche Molecular Systems, Inc.; Barbara Brady, Roche Molecular Systems, Inc.; Sung Lee, Roche Molecular Systems, Inc.

Key Words: cancer, classification, clinical, microarray, protocol, visualization

It is possible to use microarray experiments to classify cancer and other diseases. The protocol including assay of microarray experiments can change during research and development. These changes may influence the signals of microarray experiments and hence influence the classification results. We propose a method called IASCCA (Impact of Additive Signal Changes on Classification Accuracy) to use limited number of cell line experiments and existing clinical training data set to generate many thousands or millions virtual data to simulate the signal changes of the clinical data due to protocol changes and assess the impact of protocol changes on the classification accuracy. We also developed a visualization tool, projection plot, which can be used to show the results of IASCCA.

529 GIS, Map Accuracy, and Statistics ●

Section on Statistics and the Environment

Thursday, August 2, 10:30 am–12:20 pm

GIS, Map Accuracy, and Statistics

❖ Michael Goodchild, University of California, Santa Barbara, 5707 Ellison Hall, Santa Barbara, CA 93106-4060, *good@geog.ucsb.edu*

The widespread adoption of geographic information systems has renewed interest in models of map accuracy. This presentation reviews literature on the topic, and summarizes recent advances within the theoretical framework of geostatistics. The concept of measurement-based GIS is presented as a way of addressing the fundamental problems associated with the traditional coordinate-based approach. The case of area-class maps remains an outstanding problem, despite the overwhelming importance of this map type in GIS applications. The state of knowledge in modeling error in area-class maps is reviewed, and a set of six tests is proposed against which any model should be measured. It is shown that only one model meets all six tests, and that the model is in practice greatly overspecified.

Error Models in Geographic Information Systems Vector Data Using Bayesian Methods

❖Kimberly Love, Virginia Polytechnic Institute and State University, Department of Statistics, 1307 University City Blvd, Apt 3, Blacksburg, VA 24060, *krlove@vt.edu*; Eric P. Smith, Virginia Polytechnic Institute and State University; Stephen Prisley, Virginia Polytechnic Institute and State University; Keying Ye, University of Texas at San Antonio

Key Words: error, GIS, vector data, Bayesian

Existing models for vector data error in the field of GIS center on a bivariate normal model for points, and this normal model extends to line segments and polygons. In this talk we propose to incorporate Bayesian methodology into this existing model, which presents multiple advantages over existing methods. Bayesian methods allow for the incorporation of expert and historical knowledge, and reduce the number of observations required to perform an accurate analysis. This is essential to the field of GIS, where multiple observations are rare and outside knowledge is often very informative. We will explore this addition and provide some examples.

Using the Linearity of Map Features in Statistical Modeling of Map Positional Error

❖Jarrett Barber, University of Wyoming, Department of Statistics, Department of Statistics, Dept 3332 1, Laramie, WY 82071, *jbarber8@uwyo.edu*

Key Words: bayesian, GIS, linear features, maps, positional error

Map positional error refers to the difference between a feature's coordinate pair on a map and the corresponding true, unknown coordinate pair. In a geographic information system (GIS), this error is propagated through all operations that are functions of position, so that lengths, areas, etc., are uncertain. Often, a map's metadata provides a nominal statement on the positional error of a map, and such information has frequently been used to study the propagation of error through such operations. This talk will present a Bayesian statistical model for map positional error, incorporating positional error metadata as prior information, along with map coordinates, and, in particular, the information contained in the linearity of features. We demonstrate that using information in the linearity of features can greatly improve the precision of true location predictions.

Modeling the Probability Distribution of Positional Errors Incurred by Residential Address Geocoding for a Rural Health Study

❖Dale Zimmerman, The University of Iowa, 241 Schaeffer Hall, Iowa City, IA 52242, *dale-zimmerman@uiowa.edu*

Key Words: Geocode, Spatial epidemiology, Positional accuracy, Location uncertainty, Mixture distributions

Automated geocoding, which attempts to match each subject's address to an address-ranged street segment georeferenced within a streetline database and then interpolates the position of the address along that segment, is an important data assimilation step in many geographic public health studies. Unfortunately, geocoding may result in many positional errors. Our study sought to model the probability distribution of positional errors associated with automated geocoding (both 100%-matched and 60%-matched) and E911 geocoding for data from a rural health study in Iowa. We found mixtures of bivariate t distributions with few components to be flexible enough to fit these data, yet parsimonious enough to be feasible for nascent applications of measurement-error methodology to spatial epidemiology. We conjecture that bivariate t mixtures will generally be useful for geocoding errors.

530 Survival Analysis in Medical Device Studies ● ✪

Biopharmaceutical Section, WNAR, ENAR, Biometrics Section

Thursday, August 2, 10:30 am–12:20 pm

Propensity Score–Adjusted Survival Analysis

❖Nathan Carter, Boston Scientific Corporation, 1400 Hamline Avenue, South Town Squate MS 9-315, St. Paul, MN 55112, *nathan.carter@bsci.com*; Ed McMullen, Boston Scientific Corporation; Jennifer Mischke, Boston Scientific Corporation

Key Words: Propensity Scores, Survival Analysis, Historical Control, Simulation

Propensity score analysis allows adjustment for differences in baseline characteristics between two study populations. This analysis is commonly used in single arm clinical trials with a historical control. An example of this would be a stent trial, which compares the performance of a new stent to a historical control of an FDA previously approved stent. This presentation will examine the effect that propensity score adjustment has on the survival analysis conducted for a historical control stent trial. Additionally, simulations will be done to examine the possible effects propensity adjustment could have on survival analysis.

Time-to-Event Analysis Considerations in the Medical Device Industry

❖Helen Chmiel, Boston Scientific Corporation, 5077 Needham Ave NE, St Michael, MN 55376, *hmchmiel@aol.com*

Key Words: Time to Event Analysis, Survival Analysis, Proportional Hazards, Time dependent covariates

In many multi-center, randomized, controlled clinical trials in the device industry, the primary endpoint may be a time to event for several important events such as myocardial infarction, revascularization, death, or cardiovascular death. There may be a need to analyze these data for intervals that begin at a time other than Day 0 as well as multiple occurrences of an event within the interval being considered. Correctly addressing these statistical issues is critical for proper interpretation of treatment effect in this setting. This investigation is to consider different model building strategies which will look at assessing a fitted Proportional Hazards model with possible selection of model covariates that would be appropriately used for this analysis.

An Application of Propensity Score Analysis in Medical Device Clinical Studies

❖Wenji Pu, Medtronic Inc., 3415 Chandler Road, Shoreview, 55126, *wenji.pu@medtronic.com*

Key Words: Medical devices, propensity score, proportional hazards

Observational studies are common in medical device clinical studies. In those studies, patients are not randomly selected for each treatment, so the benefit from randomized studies such as balance in covariates between two treatment groups no longer exists, and treatment groups may not be comparable at baseline. Propensity score analysis can be used for adjusting for imbalance in observed covariates, and improving comparability between treatment groups. An application of the propensity methodology in a medical device study is presented with the objective to compare survival curves between ICD and non-ICD patients. To balance significant differences at baseline between the two groups, a propensity score analysis is employed under Cox's proportional hazard framework.

Assessment of Goodness-of-Fit of Cox Model with Multiple Predictors in Drug Eluting Stent Trial Efficacy Evaluation

Frank Zhou, Boston Scientific Corporation; Yukiko Imai, Boston Scientific Corporation; ❖Hong Wang, Boston Scientific Corporation, 100 Boston Scientific Way, M1 Dept of Biostatistics Clinical Science, Marlborough, MA 01752, *Hong.Wang@bsci.com*

Key Words: COX Model, Goodness-of-fit, Residuals, Drug-Eluting Stent

In this paper, we examined goodness-of-fit of Cox regression model assessing co-morbidities, cardiac risk factor, and lesion characteristics at baseline and post-procedure as predictors for the clinical outcome of target lesion revascularization (TLR) in the randomized trials of drug eluting stent. The predictors of TLR based on fitted main effect Cox model were evaluated for their goodness-of-fit based on standard error, 95% confidence interval, p-value and residual from each parameter. The Cook's distance statistic plots are used to demonstrate the effect on TLR by each particular subject. The overall goodness-of-fit tests based on grouping quartiles of risk have been applied to the model for its interpretation. Inference on treatment effect before and after the goodness-of-fit assessment were compared.

531 Application of Causal Inference in Health Sciences ✪

ENAR, Biometrics Section, WNAR, Section on Health Policy Statistics, Section on Statistics in Epidemiology

Thursday, August 2, 10:30 am–12:20 pm

A Structural Equation Model for fMRI Data with Mixed Functional and Discrete Responses

❖Wesley Thompson, University of Pittsburgh, 6476 Monitor Street, Pittsburgh, PA 15217, *wesleyt@pitt.edu*

Key Words: fMRI experiment, functional linear models, SEM, effective connectivity

In fMRI effective connectivity analyses, a common question is how two or more regions of interest (ROIs) within the brain interact with each other to process a given experimental task; it may be of interest to determine how differing levels of activation or connectivity in the ROIs affects success or failure in the completion of the task. Conversely, how a subject performs on the task may in turn affect future activation levels in the ROIs. In this talk I develop a structural equation model where some of the outcomes are discrete (success/failure) and some of the outcomes are functional (BOLD activation levels in ROIs over time). This model is applied to data obtained from an event-related fMRI working memory experiment designed to determine brain correlates of cognitive control and emotional reactivity to errors in performing tasks.

Causal Approaches to Mediation Analyses

❖Thomas R. Tenhave, University of Pennsylvania, 607 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, *ttenhave@cceb.med.upenn.edu*; Jennifer Faerber, University of Pennsylvania; Marshall Joffe, University of Pennsylvania

In the context of two randomized psychiatry trials of behavioral interventions, we present a linear rank preserving model approach for analyzing the mediation of a randomized baseline intervention by a post-randomization factor on a follow-up outcome. Unlike standard mediation analyses, our approach does not assume that the mediating factor is randomly assigned to individuals (a form of sequential ignorability). However, there is a tradeoff with other assumptions. Consistent estimation of causal mediation effects without sequential ignorability employs weights under the G-estimation approach that are optimal in terms of semi-parametric efficiency but under sequential ignorability. In this context, we will present analyses of mediation of behavioral intervention effects on reducing depression outcomes by medication and therapy factors in the psychiatry trials.

Confounders in Assessment of Haplotypes for Lung Cancer Risk

❖Roger Day, University of Pittsburgh, 310 Cancer Pavillion, 5150 Center Avenue, Pittsburgh, PA 15232, *day@upci.pitt.edu*

Key Words: causal models, haplotype, cancer, epidemiology, confounding

In studying genetically-caused susceptibility to lung cancer, potential confounders must be handled with more care than is commonly exercised in regression analyses of observational studies. Each of the following factors—family history of lung cancer, age, race, and smoking—presents unique challenges and different consequences for adjustment. Theories of causal analysis provide a framework for assessing causal effects in distinction to associations, in the presence of other causes and associations. Causal diagrams can help one develop and critique analysis strategies. In some circumstances, the causal calculus of J. Pearl provides consistent estimators for causal relationships, given a reasonably complete causal diagram. This is often true even with the presence of unmeasured factors in the diagram. Advantages and limitations of this theory are examined in the lung cancer study.

A Longitudinal Mediational Model To Assess the Indirect Effect of Adolescent-Physical Activity in Young Adult-Physical Activity Between Males and Females

❖Kang Sun, University of Pittsburgh, Department of Biostatistics, Graduate School of Public Health, 210 Handler St Apt2, Pittsburgh, PA 15203, *kas34@pitt.edu*; Vincent C. Arena, University of Pittsburgh

Mediational analysis is used to explain "how" a predictor affects an outcome through an intervening variable called a mediator. A mediational analysis can be based on longitudinal modeling. We describe one such analysis using data from the Adolescent Injury Control Study where 1245 adolescents were followed from adolescence (Phase I) to young adulthood (Phase II). Physical activity hours and several risk factors were surveyed annually. A longitudinal mediation model is used to examine the role of adolescents-physical activity (mediator) in the relationship between gender (predictor) and the physical activity in young adults (outcome). The initial status and changes in the mediator and the outcome are included in one multivariate mediational model. The indirect effect of the mediator is assessed by the latent growth curve modeling techniques.

532 Methodological Issues in Engineering Applications of Computer Models: A SAMSI Program ● ✪

Section on Bayesian Statistical Science
Thursday, August 2, 10:30 am–12:20 pm

Resampling for Inverse Problems Involving High-Dimensional Computer Simulators

❖ Herbert Lee, University of California, Santa Cruz, School of Engineering, 1156 High Street, Santa Cruz, CA 95064, *herbie@ams.ucsc.edu*; Matthew Taddy, University of California, Santa Cruz; Bruno Sanso, University of California, Santa Cruz

Key Words: Bayesian statistics, Importance Sampling, Gaussian process, soil permeability

The two most common Bayesian approaches to computer model inverse problems are surrogate models or direct MCMC using the simulator. However, in some cases the simulator may have already been run on a large number of input cases, and may no longer be available for additional runs. When tens of thousands or more runs are available, standard surrogate models such as Gaussian processes are too computationally expensive to fit. When no additional runs are available, direct MCMC is not possible. Thus we have developed a sampling importance resampling algorithm that is applicable in such cases. We demonstrate our approach on a hydrology problem where we want to draw posterior inference for an unobserved high-dimensional spatial permeability field using a million runs from a simulator of water flow.

A Bayesian Analysis of the Thermal Challenge Problem

❖ Fei Liu, Duke University, Institute of Statistics and Decision Sciences, Box 90251, Durham, NC 27708, *fei@stat.duke.edu*; Susie Bayarri, University of Valencia; James Berger, Duke University; Rui Paulo, ISEG-Technical University of Lisbon; Jerry Sacks, National Institute of Statistical Sciences

Key Words: Bayesian analysis, Computer model validation, Gaussian stochastic process, Thermal computer model

A major question for the application of computer models is “Does the computer model adequately represent reality?” Viewing the computer models as a potentially biased representation of reality, Bayarri et al.(2005) develop the Simulator Assessment and Validation Engine (SAVE) method as a general framework for answering this question. In this paper, we apply the SAVE method to the challenge problem: a thermal computer model designed for certain devices, and develop a statement of confidence that the devices can be applied in intended situations.

Bayesian Kalman Filter for Emulation of Complex Computer Models

❖ Gentry White, North Carolina State University, 2501 Founders Drive, 210C Patterson Hall, Raleigh, NC 27695-8203, *white@stat.ncsu.edu*; Peter Reichert, Swiss Federal Institute of Aquatic Science and Technology

Key Words: Bayesian, Emulator, State-space Model, Kalman FIlter, Dynamic Systems

Non-linear computer models can be used to simulate complex dynamic systems. The expense of running these models limits exploration of overall system behavior. In this cases a simpler linearized computer model called an emulator is desirable to explore system behavior under a variety of inputs. A good emulator should include the relevant physics for the system, perfectly interpolate the observed data and provide reasonable estimates of the complex model output when there are no observations. In the case of modeling dynamic systems we apply a state-space model, using a Kalman filter in the context of Bayesian estimation in order to construct an emulator. This model is a Gaussian Process and thus has good emulator properties. The model developed here is shown to provide a reasonable emulator for a real world dataset concerning a hydrological ground water runoff model.

Determining the Pareto Frontier for Multiobjective Optimization of Black Box Functions

Dianne Bautista, The Ohio State University; ❖ Thomas Santner, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Cockins Hall, Room 404, Columbus, OH 43210, *tjs@stat.ohio-state.edu*

Key Words: Multiobjective optimization, Pareto Frontier

Most engineering design efforts involve the optimization of multiple price/performance or other competing objectives. The Pareto frontier of a set of functions specifying these multiple objectives for a given application, consists of those engineering designs that cannot be simultaneously improved for all objectives. This talk describes a method of determining the Pareto frontier when the competing functions correspond to expensive-to-calculate computer codes. Methods for identifying the Pareto Frontier in the computer experiment setting will be discussed.

Aleatory and Epistemic Uncertainty Quantification for Engineering Applications

❖ Laura Swiler, Sandia National Laboratories, Sandia Labs PO Box 5800, MS 1318, Albuquerque, NM 87185, *lpswile@sandia.gov*; Anthony Giunta, Sandia National Laboratories

Key Words: uncertainty quantification, engineering models, aleatory, epistemic

Most computer models for engineering applications are developed to help assess a design or regulatory requirement. As part of this task, the capability to quantify the impact of variability and uncertainty in the decision context is critical. The requirement is often stated as: the probability that some system response quantity exceeds a threshold value is less than some required probability. This presentation will provide an outline and comparison of methods that are used for analyzing and propagating aleatory and epistemic uncertainties. The methods are all available in a software tool called DAKOTA. We will specifically focus on four classes of methods: Latin Hypercube sampling, Dempster-Shafer theory of evidence, “second-order” probability analysis, and analytic reliability methods. Examples of each of the methods as applied to a simple engineering model will be provided.

533 Nonparametric and Semiparametric Methods for Data with Measurement Errors ●

Section on Nonparametric Statistics, Section on Statistics in Epidemiology
Thursday, August 2, 10:30 am–12:20 pm

Bandwidth Selection for Weighted Kernel Density Estimation

❖ Bin Wang, University of South Alabama, ILB 325 University Blvd, Mobile, AL 36688, *bwang@jaguar1.usouthal.edu*

Key Words: informative censoring, nonparametric, bandwidth selection, weighted kernel density estimate, missing data

Censoring is common in survival data. For noninformatively censored data, a weighted kernel density estimator can be used to estimate the density of the survival times. However, bandwidth selection could be a problem by reweighting the kernels. In this paper, we proposed to justify the bias cause by noninformative/informative censoring by a weighted kernel method and the bandwidth selection will be discussed. The performance of the new method will be illustrated via Monte Carlo studies.

Multivariate Failure Time Data Modeling with Instrumental Variables

❖ Zhaozhi Fan, Memorial University of Newfoundland, 22 Brixham Cres, Torbay, NF A1K 1N7 Canada, *zhaozhi@math.mun.ca*; Xiaofeng Wang, The Cleveland Clinic

Key Words: marginal hazards model, multivariate failure time, measurement error, semiparametric, intrumental variables

We consider marginal hazards model of multivariate failure times with covariates measured with error and with instrumental variables available. We extend the kernel smoothing procedure of Zhou and Wang (2000) to correlated failure time data. We estimate the marginal partial likelihood function with a semiparametric procedure. We present asymptotic properties of the induced estimator and provide some simulation results.

Estimating the Distribution Function with Measurement Error in a Complicated Survey

❖ Liang Li, The Cleveland Clinic, 9500 Euclid Ave Wb4, Cleveland, OH 44195, *lil2@ccf.org*; Tom Greene, The Cleveland Clinic

Key Words: chronic renal disease, density estimation, iSIMEX, measurement error, NHANES, SIMEX

The simulation extrapolation (SIMEX) method can be used to estimate the cumulative density function of a random variable that is unobservable and measured with additive error (Stefanski and Bay, 1996). Motivated by a study using the NHANES survey, we extend the SIMEX methodology by adding iterative bootstrap bias reduction steps. This iterative SIMEX (iSIMEX) method is shown to have less bias and mean square error than the SIMEX in our problem, and the resulting cumulative density function is monotone and constrained between 0 and 1. We apply the proposed methodology to estimating the population prevalence of people with various levels of chronic renal disease, and population distribution of kidney function in the US.

A New Approach for Measurement Errors in Linear Models

❖ Guan Xing, Bristol-Myers Squibb Company, 180 Franklin Corner Rd, Apt L9, Lawrenceville, NJ 08648, *guan.xing@bms.com*

Key Words: measurement error, mixture model

Measurement errors happens often in real research. We see the measurement error problem as a mixture model with unknown number of components and use the approach proposed by Xing and Sunil (2006) to do model selection. The performance is demonstrated with simulation study and real dataset.

534 Statistical Analysis of Neuronal Data ●

IMS

Thursday, August 2, 10:30 am–12:20 pm

Using Point Process Models To Describe Oscillatory Spiking in the Subthalamic Nucleus of Parkinson's Patients

❖ Uri Eden, Boston University, 111 Cummington St, Boston, MA 02215, *tzvi@bu.edu*; Emery N. Brown, Massachusetts General Hospital; Emad Eskandar, Massachusetts General Hospital

Key Words: Neural Data, Point Processes, GLM

Abnormal oscillatory firing patterns of neurons in the subthalamic nucleus (STN) of patients with Parkinson's disease (PD) have been postulated to play a role in the pathogenesis of motor deficits. We characterized the firing properties of these neurons by constructing GLM point process intensity models using spline basis functions that relate the spiking of each neuron to movement variables and the neuron's past firing history, both at short and long time scales. By calculating maximum likelihood estimators for all of the parameters and their significance levels, we were able to describe the relative propensity of aberrant STN spiking in terms of factors associated with reaching movements, with intrinsic properties of the neurons, and factors that may be related to disregulated network dynamics and to determine their relative importance in describing the observed data.

Modeling and Decoding Multineuronal Responses in the Primate Retina

❖ Liam Paninski, Columbia University, Department of Statistics, 1255 Amsterdam Ave, New York, NY 10027, *liam@stat.columbia.edu*

Key Words: neuroscience, retina, point process, GLM, log-concavity, network modeling

A key challenge in neuroscience is to understand how large networks of neurons collectively encode information. It has recently become possible to record simultaneously from complete mosaics of ~100 ON and OFF parasol cells over a 4x8 degree region of peripheral retina, providing the opportunity to better understand the message the eye sends the brain. We present a new generalized linear multivariate point process model of the visual response properties of networks of these retinal ganglion cells. We discuss model fitting and validation (challenging due to the high dimensionality of the data, but computationally tractable due to the concavity of the loglikelihood), demonstrate the model's accuracy in predicting retinal responses to novel visual stimuli, and describe Bayesian techniques for decoding the network responses. With J. Pillow, E. J. Chichilnisky, E. Simoncelli, and J. Shlens.

Real-Time Decoding of Nonstationary Motor Cortical Activity

❖ Wei Wu, Florida State University, 1015 W Call St, Dept of Statistics, Tallahassee, FL 32306-4330, *wwu@stat.fsu.edu*

Key Words: decoding, real-time, adaptive, motor cortex, nonstationarity, Kalman filter

Neural decoding has played a key role in recent advances in motor cortical Brain-Machine Interfaces. A number of practical algorithms have been proposed to predict continuous movement in an on-line circumstance using recordings from a chronically implanted multi-electrode micro-ar-

ray in the motor cortex. Though effective, those models were based on a strong assumption that the neural signal sequences are a stationary process. Recent work, however, indicates that the motor system significantly varies over time. To characterize the dynamic relationship between neural signals and hand kinematics, here we develop an adaptive approach based on the well-known Kalman filter. This new approach recursively updates encoding model in an on-line fashion. Experimental results show that it decodes more accurately than a non-adaptive Kalman filter, while remains real-time efficiency.

Jitter Methods for Investigating the Time Scale of Dependencies in Neuronal Firing Patterns

❖Matthew Harrison, Carnegie Mellon University, Baker Hall 132C, Pittsburgh, PA 15213, *mtharris@cmu.edu*; Asohan Amarasingham, Rutgers University; Stuart Geman, Brown University

Key Words: point process, time scale, dependencies, correlations, nonparametric

Detecting dependencies in the spatio-temporal pattern of neuronal spiking activity is challenging, in part because many different types of dependencies are confounded in the data. Classical permutation tests can often be used to distinguish time-locked from non-time-locked dependencies. These methods, however, do not easily disambiguate interactions on multiple time scales and do not extend to situations with single or non-repeatable samples. We have been developing techniques, called jitter methods, designed to investigate the time scale of arbitrary spatio-temporal dependencies in neuronal spiking activity. Jitter methods are analogous to classical permutation tests except that the permutations are over spike times, instead of trial labels. Jitter methods perform well in a variety of settings where dependencies on multiple time scales are confounded.

Variable Selection and Classification Using Computed Tomography (CT) Medical Image Data

❖Hyun Kim, University of California, Los Angeles, 924 Westwood Blvd ste # 650, Department of Biostatistics, Los Angeles, CA 90024, *gracekim@mednet.ucla.edu*; Gang Li, University of California, Los Angeles; David W. Gertson, University of California, Los Angeles; Robert Ochs, University of California, Los Angeles; Matthew S. Brown, University of California, Los Angeles; Jonathan Goldin, University of California, Los Angeles

Key Words: variable selection, classification, Computed Tomography, high dimension, texture, image

PURPOSE: CT data is high dimensional and spatially correlated. The purpose of this study is to evaluate the sensitivity of two classification methods, namely, multinomial logistic regressions using backward selection and non-concave penalized likelihood feature selection using support vector machine classification for lung disease using texture features. METHOD: Each method was evaluated with and without clustering. CONCLUSIONS: For the multinomial model the classification accuracy was 91.4% and 91.1% with and without subject cluster respectively. The SVM classification rate was 91.7%. 40 features were commonly selected. In textural classification of lung diseases, accounting for possible intra-subject dependencies in the training set does not affect the resulting classification on a separate test set.

535 Estimating the Incidence, Effect, and Cost of New Treatments for Health Outcomes ● ✪

Section on Health Policy Statistics, ENAR, Biometrics Section, WNAR

Thursday, August 2, 10:30 am–12:20 pm

How Do You Argue Against the Use of a High-Cost Medical Treatment When a 'Quality-of-Life' Measure Is Not Available?

❖Andrea McAllister, Highmark Inc., 120 Fifth Avenue, Suite P7205, Pittsburgh, PA 15222, *andrea.mcallister@highmark.com*

Key Words: high-cost, treatment, healthcare, insurance

A pre/post total healthcare cost analysis is performed on all patients receiving an initial newly introduced asthma treatment. There were slight increases in per patient per month (PPPM) costs for medical (non-study treatment) claims and outpatient ER visits, and larger decreases in asthma pharmacy. However, the combined impact of these changes is minimal compared to the increase in medical costs related to the study treatment. What is missing from this study is a measure of the impact that this new asthma treatment has on the patient's quality of life. Even though there was a small increase in Outpatient ER visits, there was a decrease in inpatient stays which can be assumed to significantly impact a patient's quality of life. What about the impact of the drug treatment itself? Unless we directly contact these patients and survey them personally, we can't.

Estimation of Incidence of First-Time-Ever Treatment, Prevalence of Paused Treatment, and Prevalence of Active Treatment from Pharmacoepidemiologic Claims Data

❖Henrik Stovring, University of Southern Denmark, Research Unit of General Practice, JB Winslows Vej 9A, Odense, 5000 Denmark, *hstovring@health.sdu.dk*

Key Words: Incidence, Prevalence, Paused treatment, Pharmacoepidemiology, Censoring, Waiting time distribution

In pharmacoepidemiology, estimation of incidence and prevalence of drug use is of key interest. The basic epidemiologic methods for incidence and prevalence do however impose a chronic treatment assumption which is unrealistic for most medications. Hence we suggest to extend the waiting time distribution model (time from a fixed date to first subsequent treatment event for each patient) to three compartments, corresponding to no previous, paused, or current treatment, respectively. As the two first compartments can not be distinguished in the waiting time distribution alone, we incorporate information from the interarrival distribution for time between treatment events. This transformation is non-trivial as it introduces informative censoring into the waiting time distribution, and we therefore consider how censoring can be modeled to allow valid estimation of the interest parameters.

Mixed Treatment Comparison of Migraine Headache Treatment: Indirect Comparison and Evidence Consistency

❖ Rongwei (Rochelle) Fu, Oregon Health & Science University, Dept of Public Health And Preventive Medicine, 3181 S.W. Sam Jackson Park Rd., Portland, OR 97239-3098, *fur@ohsu.edu*; Kimberly Peterson, Oregon Health & Science University

Key Words: Mixed treatment comparison, evidence inconsistency, meta-analysis, indirect comparison, migraine

In drug class systematic review, effectiveness and safety comparisons between drugs in the same class are made and the results are often used to inform public policy and decision making. It is common that head-to-head trials are not available to compare some drugs and mixed treatment comparison enables comparison of treatments effects that are not directly observed, or strengths direct estimates with indirect estimates. Evidence consistency is one of the requirements that make mixed treatment comparison. Triptan is a class of drug treating migraine headache and there are seven triptans available in U. S. In this study, we will perform a Mixed Treatment Comparison of about 45 triptan trials and compare the current methods on mixed treatment comparison and evidence consistency assessment. The effects of encapsulation on effectiveness is also planned to be investigated.

Effect of Intravenous Insulin Infusion on Coronary Artery Disease

Samer Ellahham, Innovative Medical Institute/Paragon Cardiovascular Foundation; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University; ❖ Rima Allaham, Paragon Cardiovascular Foundation, , *samer1@ix.netcom.com*

Key Words: clinical manifestations, CAD

Coronary Artery Disease (CAD) is the leading cause of death in adults in the United States. National health and nutrition surveys suggest that at around 14.0 million people in the United States have CAD. CAD can cause myocardial ischemia, which can be clinically stable or acute coronary syndrome. Atherosclerosis is an inflammatory state and the inflammatory markers; hs-CRP and SAA have been shown to be predictive of clinical manifestations and complications of coronary artery disease. Pro-atherogenic stimuli increase generation of reactive oxygen species and transcription factors. These phenomena occurring in the coronary artery present clinically as acute coronary syndrome. Insulin infusion have been demonstrated in clinical trials of acute myocardial infraction in diabetics and non-diabetics to be safe and efficacious.

Update on Cardiometabolic Risk Reduction: Role of the Endocannabinoid System

❖ Samer Ellahham, Innovative Medical Institute/Paragon Cardiovascular Foundation, 23044 Winged Elm Drive, Clarksburg, MD 20871, *samer1@ix.netcom.com*; Yasmin H. Said, George Mason University; Edward Wegman, George Mason University

Key Words: clinical effectiveness, rimonabant

Obesity is the most common nutritional disorder in western industrialized countries and arises from the accumulation of excess fat in the body from over consumption of fatty foods. The prevalence of obesity and diabetes is increasing dramatically in the Sates and worldwide. Obesity is a major risk factor for diabetes, cardiovascular disease, and metabolic syndrome. The metabolic syndrome is a precursor to cardiovascular disease and diabetes. The endocannabinoid system controls food intake via both central and peripheral mechanisms. The cannabinoid receptor antagonist effect has a broad impact on metabolism and is associated with several sites of action and mechanisms. The administration of CB1 antagonists has several clinical implications. Blockade of the endocannabinoid system has therapeutic potential in the management of obesity and cardiometabolic risk.

536 Issues Related to Power and Sample Size Calculation ● ✪

Biometrics Section, Section on Teaching Statistics in the Health Sciences

Thursday, August 2, 10:30 am–12:20 pm

Power and Sample Size Estimation for the Wilcoxon Rank Sum Test

❖ Bernard Rosner, Harvard Medical School, 181 Longwood Avenue, Channing Laboratory, Boston, MA 02115, *bernard.rosner@channing.harvard.edu*; Robert Glynn, Brigham and Women's Hospital

Key Words: Mann-Whitney U Test, receiver operating characteristic curve, probit transformation, power

The Wilcoxon Mann-Whitney (WMW) Test is commonly used in non-parametric two group comparisons when normality assumptions are questionable. There has been some previous work on estimating power for this procedure based on specific underlying distributions (Pettitt and Siskind). In this paper, we present an approach for estimating type II error which is applicable to any continuous distribution. We apply these results to obtaining standard errors of the area under the receiver operating characteristic curve (AUROC) for risk-prediction rules and for comparing AUROC between competing risk prediction rules applied to the same data set. These results can be implemented using SAS-callable functions to evaluate the bivariate normal integral and are less computer-intensive than previously published approaches of DeLong, et al. (1988) to address this problem.

Sample Size Calculations in Simple Linear Regression: A New Approach

❖ Marepalli B. Rao, University of Cincinnati, Center for Genome Information, 3223 Eden Avenue, Cincinnati, OH 45267, *marepalli.rao@uc.edu*; Mohammed K. Alam, Kendle International Inc.; Subramanyam Kasala, The University of North Carolina at Wilmington; Thomas Bradstreet, Merck & Co., Inc.

Key Words: Linear Regression, Sample Size, Size, Power

In the context of a simple linear regression, a typical problem is to test the null hypothesis that the slope parameter is zero against the alternative that it is not equal to zero. If we have data on the response variable and the covariate, we can calculate the power of the standard test used at any specified alternative value of the slope parameter. For given size, power, and the alternative value of the slope parameter, if we want to calculate sample size required, any sample size software demands knowledge of the sum of squares of the covariate values. We will not have this information. In this talk, we will outline how to overcome this problem. In the course of this investigation, we came across a new distribution, which is of pedagogic interest by itself.

Why Family-Based Association Makes Sense in a Genome-Wide World: A Comparison of Study Power

❖ David Fardo, Harvard University, 655 Huntington Ave, SPH2 Bldg II 4th Floor, Boston, MA 02115, *dfardo@hsph.harvard.edu*; Christoph Lange, Harvard School of Public Health

Key Words: genome-wide association, joint analysis, multi-stage testing, multiple comparisons, power, statistical genetics

Genetic association studies can be broken into two basic categories: population-based studies in which unrelated subjects are recruited and family-based studies which rely on related subjects. The distinguishing characteristics of these two fundamental approaches have been addressed in a number of papers (Laird and Lange, 2006 and Hopper et al., 2005 among others). Recent technological and computational advances have made genome-wide association studies a reality (Herbert et al, 2006; Klein et al., 2005; Cheung et al., 2005) and have also made it necessary to consider the implications of employing different designs. Although these genome-wide studies are now feasible from a cost perspective, it is of utmost importance that they be designed to be cost-efficient, powerful and robust to modeling assumptions. We investigate here differences in statistical power between popular approaches.

Sample Size Calculations for Clinical Trials with Rate of Decline on a Quantitative Trait as the Endpoint

❖ Steven Edland, University of California, San Diego, 9500 Gilman Drive MC0949, La Jolla, CA 92093-0949, *sedland@ucsd.edu*

Key Words: Clinical Trial, Sample Size, Phase III, Power, Compound symmetry, Pilot Data, Compound symmetry, Pilot Data

While estimates of marginal model parameters by gee are robust to misspecification of the covariance structure, we caution that power calculation formulas that depend on covariance structure assumptions are not. We illustrate this mathematically and empirically using Monte Carlo simulations informed by data from a recently completed clinical trial of a treatment to slow the progression of Alzheimer's disease. We also caution that covariance estimates from shorter term pilot studies may not be representative of the covariance that will be observed in the clinical trials to be powered, because the covariance often depends on the length of follow-up and interval between observations. We derive an alternative power formula that can be used to power clinical trials of arbitrary length and interval between follow-up that does not require explicit assumptions about the covariance structure.

Mixed-Effects Logistic Model for Association Analysis

❖ Sanjay Shete, The University of Texas M.D. Anderson Cancer Center, 1155 Pressler Blvd CPB43628, Houston, TX 77030, *sshete@mdanderson.org*

Key Words: Mixed-Effects model, case-control study, association, covariates, power

An association study to identify possible causal genes is a popular approach. However, in association studies cases and controls are assumed to be genetically unrelated. In certain situations, genetic data may be available on other affected family members. Choosing a single affected individual per family is statistically inefficient and leads to a loss of power. On the other hand using affected family members and unrelated normal controls directly leads to false-positive results. We propose a new approach using mixed-model logistic regression, in which associations are performed using family members and unrelated controls. Extensive simulation studies showed that our approach can effectively control the type-I error probability and has higher power than likelihood based methods. We applied this method to correlate mutagen sensitivity and genes involved in NER pathway using a twin study.

Covariate-Modulated False Discovery Rates

❖ Egil Ferkingstad, University of Oslo, Department of Biostatistics, PO Box 1122 Blindern, Oslo, N-0317 Norway, *egil.ferkingstad@medisin.uio.no*; Arnoldo Frigessi, University of Oslo; Gudmar Thorleifsson, deCODE genetics; Augustine Kong, deCODE genetics

Key Words: multiple testing, bioinformatics, data integration, genomics, false discovery rate, empirical Bayes

In an empirical Bayes setting, the local false discovery rate (local FDR) is defined as the posterior probability that the null hypothesis is true given data. We extend this methodology further and introduce the covariate-modulated false discovery rate (cmFDR), useful when an additional covariate is available that influences the probability of each null hypothesis being true. cmFDR measures the posterior significance of each test conditionally on the covariate and the data, leading to greater power. The cmFDR uses covariate-based prior information to produce a list of significant hypotheses which differs in length and order from the list obtained by the local FDR. We estimate the cmFDR with MCMC for an approximate model on p-values. The new method is applied to expression quantitative trait loci (eQTL) data, and to gene expressions modulated by copy number alterations in breast cancer.

Regression-Based Association Approach Using Genetic Similarity for Genomewide Association Scans

❖ Jung-Ying Tzeng, North Carolina State University, Department of Statistics, Campus Box 7566, Raleigh, NC 27695, *jytzeng@stat.ncsu.edu*; Sheng-Mao Chang, North Carolina State University; Duncan C. Thomas, University of Southern California; Marie Davidian, North Carolina State University

Key Words: multi-marker analysis, genetic association, haplotype analysis, genome-wide association studies

Genomewide association scans (GWAS) become a new tool to identify genetic variants underlying complex traits. While multimarker analyses have been frequently used in association studies, most statistical analyses of GWAS still focus on single-SNP analyses. One concern for multimarker analyses is its large degrees of freedom (df) for capturing genetic diversity. As the power of a single test is limited by the large df, the overall performance of a GWAS is further diminished by multiple testing. Here we consider a regression approach for testing association using genetic similarity. The method is inspired by the Haseman-Elston regression for linkage analysis and is connected with the ordinary regression that treats haplotypes as covariates under hierarchical modeling framework. Via simulation we assess its performance and demonstrate its validity and power in testing genetic association.

537 Issues Related to Clinical Trials ● ✪

Biometrics Section, Biopharmaceutical Section

Thursday, August 2, 10:30 am–12:20 pm

Data Monitoring in Clinical Trials Using Prediction

❖ Scott Evans, Harvard University, Harvard School of Public Health, 651 Huntington Ave, Boston, MA 02115, *evans@sdac.harvard.edu*; Lingling Li, Harvard School of Public Health; Lee-Jen Wei, Harvard University

Key Words: data monitoring, prediction, predicted intervals, interim analyses, clinical trials

Clinical trials (CTs) are often monitored for efficacy or futility. Several methods for interim monitoring of CTs have been developed. Although informative, few of these methods convey information regarding effect sizes (e.g., treatment differences) and none use prediction to convey information regarding potential effect size estimates and associated precision, with trial continuation. We propose use of prediction and specifically "predicted intervals" (PIs) as a flexible and practical tool for quantitative monitoring of CTs. PIs provide information regarding effect sizes, are invariant to study design, and provide flexibility in the decision making process. We outline construction of PIs for binary, continuous, and time-to-event endpoints and present examples of their use. PIs provide a valuable tool for Data Monitoring Committees (DMCs).

Specifications Based on Pharmacokinetic Evaluation

❖Jose Casas-Finet, MedImmune, Inc., One MedImmune Way, Gaithersburg, MD 20878, *casasj@medimmune.com*; Harry Yang, MedImmune, Inc.; Iksung Cho, MedImmune, Inc.; Mark Schenerman, MedImmune, Inc.

Key Words: Specification, PK

For a marketed drug product, the specifications are usually heavily reliant on manufacturing experience, since only a limited number of lots are used for clinical trials. This potentially could result in ranges narrower than what would have been allowed by clinical outcomes. In this paper, we propose a novel approach to evaluate the impact of the distribution of the attribute on the pharmacokinetic (PK) profile of the product lot. By modeling the theoretical relationship between the attribute and PK profile, we are able to identify a range of the attribute that produces bioequivalent drug product. The bioequivalent range, in conjunction with clinical safety data, can be used to justify specification limits that are typically wider than those derived solely from manufacturing history. The method is illustrated through a concrete example.

The Exact Distribution of the Heterogeneity Statistic in Meta-Analysis

❖Brad Biggerstaff, Centers for Disease Control and Prevention, PO Box 2087, Fort Collins, CO 80522-2087, *bbiggerstaff@cdc.gov*; Dan Jackson, University of Salford

Key Words: meta-analysis, homogeneity test, heterogeneity, West Nile virus

The exact distribution of Cochran's heterogeneity statistic Q in one-way meta-analysis is derived, and software is provided to facilitate computation. The exact Q-distribution is used to derive the exact distributions of the heterogeneity measures I^2, H^2 and H^2_M, and also of the DerSimonian-Laird estimator of the heterogeneity variance parameter. Comparisons of standard approximations are made to the exact results in applications appearing in the literature and of a meta-analysis of West Nile virus studies of clinical to sub-clinical infection ratios.

A Model-Based Assessment of Antitumor Activity for Solid Tumor Xenograft Studies

❖Jianrong Wu, St. Jude Children's Research Hospital, 332 N Lauderdale St, Memphis, TN 38016, *jianrong.wu@stjude.org*

Key Words: Antitumor activity, Exponential growth model, Tumor growth delay, Xenograft model

Three exponential tumor growth models are developed to characterize tumor growth in the mouse solid tumor xenograft studies for the Pediatric Preclinical Testing Program (PPTP) (Houghton, et al., 2006). Furthermore, a model-based antitumor activity measurement, the AUC ratio, is proposed. The typical antitumor activity measurements used in preclinical xenograft studies, such as $\log_{10}$ cell kill and tumor growth inhibition T/C could either be inapplicable or mislead the antitumor activity evaluation. The AUC ratio provides a more insightful antitumor activity evaluation throughout the entire treatment period.

Subgroup Weighted Association Testing

❖Michael LeBlanc, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M3 C102, Seattle, WA 98109, *mleblanc@fhcrc.org*

Key Words: SNP, tree-based regression, assocation studies, subgroups, logic, testing

We investigate a class of test statistics for large-scale association studies based on weighting subgroup information. Given many analyses involve testing univariate or marginal SNP associations with subject outcome, more complex relationships such as multiple gene combinations or gene-environment (or gene-treatment) interactions could reduce the power to detect the marginal associations. We propose a computationally simple panel of weighted marginal tests that can exploit such subgroup associations if they exist. The strategy modestly increases the search by focusing (or enriching) the association tests within subgroups. Extensions of the method to allow multiple environmental or clinical factors are considered and connections are made to tree-based models and logic regression models of Ruczinski, Kooperberg, and LeBlanc (2003).

Ensemble Methods for Classification of Patients for Personalized Medicine Using High-Dimensional Biomarkers

❖Hojin Moon, Food and Drug Administration, 3900 NCTR RD, HFT20, Jefferson, AR 72079, *hojin.moon@fda.hhs.gov*; Hongshik Ahn, Stony Brook University; Ralph L. Kodell, University of Arkansas for Medical Sciences; Songjoon Baek, Food and Drug Administration; Chien-Ju Lin, National Center for Toxicological Research; James J. Chen, Food and Drug Administration

Key Words: Classification tree, Cross-validation, Ensemble classifiers, Majority voting, Random partitions, Resampling

Personalized medicine is defined by the use of genomic signatures of patients in a target population for assignment of more effective therapies as well as better diagnosis and earlier interventions that might prevent or delay disease. Classification algorithms can be used for prediction of response to therapy to help individualize clinical assignment of treatment. An ensemble classifier built from the optimal number of random partitions of the feature space will be presented. This classification algorithm can overcome the problem of having fewer samples than predictors. The algorithm is applied to several published genomic data sets to classify patients into risk/benefit categories. Based on cross-validated results for several high-dimensional data sets, our algorithm is consistently one of the best classification algorithms.

Recursive Estimation Method for Predicting Residual Bladder Urine Volumes to Improve Accuracy of Timed Urine Collections

❖David Afshartous, University of Miami, School of Business room KE408, Coral Gables, FL 33124, *afshar@miami.edu*; Richard Preston, University of Miami

Key Words: repeated measures, recursive systems, residual volume, sensitivity analysis

Clinical research studies often collect data via repeated measurements of collected urine. Timed urine collections are inexpensive, noninvasive, relatively simple to conduct, and can provide valuable data on kidney function and the clearance of various electrolytes and drugs by the kidney. Unfortunately, the accuracy of timed urine collections is limited by the presence of a residual volume of urine remaining in the bladder following each timed void due to incomplete emptying of the bladder. We present an unbiased method to estimate the residual bladder volumes via a mathematical model of the bladder process. The model leverages constant concentration and constant mass principles in order to solve a system of recursive equations. We verify the model on simulated patients and also investigate the sensitivity of the model to initial value specification.

538 Regression Approaches To Analyze Discreet Data in Clinical Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Thursday, August 2, 10:30 am–12:20 pm

Application of the Zero-Inflated Negative Binomial Distribution

❖ Arlene Swern, Merck & Co., Inc., RY34B370, Rahway, NJ 07065, *arlene_swern@merck.com*; Hongwei Wang, Merck & Co., Inc.; Qian Dong, Merck & Co., Inc.; Eric C. Kleerup, David Geffen School of Medicine at UCLA

Key Words: distribution, zero inflated Poisson, count data, overdispersion, negative binomial, zero modified

Count data is typically modeled with a Poisson distribution. In real life settings however, the assumptions of this distribution are often violated because of too many or too few zero events or because of overdispersion. The zero-inflated negative binomial (ZINB) model and an extension, the zero-modified negative binomial (ZMNB) are proposed as a robust family of distributions for summarizing count data. These models are generalizations of the Poisson and negative binomial distributions, with parameters to handle the excess or deficit of zeroes and overdispersion. The amount of daily albuterol used as inhaled rescue medication is an important clinical measure of both the efficacy of asthma treatment and the extent of asthma control, but the Poisson distribution is not always appropriate. The application of the ZINB and ZMNB models to summarize this measure will be explored.

Evaluation of Risk-Benefit in New Drug Development Using a Global Event Intensity Index

❖ Jingli Song, Eli Lilly and Company, 11884 Esty Way, Carmel, IN 46033, *songji@lilly.com*; Andreas Sashegyi, Eli Lilly and Company; Steve Zheng, Eli Lilly and Company

Key Words: risk-benefit evaluation, clinical trials, Poisson regression model, weights

In a long-term clinical trial of new drug, patients may experience efficacy and safety endpoints over study period, which are usually evaluated individually. Of key interest to regulatory agencies, patients, and clinicians is whether the drug has a positive risk-benefit profile. A good evaluation of risk benefit must account for patients with more than one event, patient's exposure time, and differences in severity levels and background rates among the events. In this paper, a global event intensity index fulfilled the above requirements is proposed to evaluate the risk-benefit of a drug. A repeated measures Poisson regression model is used to analyze the weighted total number of events of a patient over the study period. Weights are predetermined based on severity levels. An example from a clinical trial of raloxifene is used to illustrate the method.

Regression Models for Multinomial Responses with an Example from PCPT

❖ Kenneth Liu, Merck & Co., Inc., UG1C46, P.O. Box 1000, North Wales, PA 19454-1099, *Kenneth_Liu@merck.com*; Alexandra Carides, Merck & Co., Inc.

Key Words: Polytomous logistic regression, nominal response, ordinal response, proportional odds assumption

Logistic regression is used to analyze a binary response variable. When the response variable has more than two levels, various forms of multinomial logistic regression can be used. We will discuss two types that permit non-proportional odds. Nominal logistic regression is a direct extension of logistic regression where each unordered response category is compared to a common reference. The unconstrained partial proportional odds model is one type of ordinal logistic regression that models an ordered response using cumulative logits. In the Prostate Cancer Prevention Trial (PCPT), patients can have high-grade, ungraded, low-grade, or no cancer. Nominal logistic regression compares each grade of cancer to a reference group such as the no cancer category. The unconstrained partial proportional odds model orders the cancer categories and compares cumulative categories.

Mixing Negative Binomial Distribution with Completely Monotonic Functions and Their Applications in Modeling Over-Dispersion Data

❖ Xiaodong Wang, sanofi-aventis, 200 Crossing Blvd, Biostatistics, P.O.Box 6890, Bridgewater, NJ 08807, *sheldon.wang@sanofi-aventis.com*; Hanxiang Peng, The University of Mississippi; Ji Zhang, sanofi-aventis

Key Words: complete monotonicity, exchangeability, negative binomial distribution

In many teratogenic and clinical studies, response binary data are often correlated within each litter or patient. Many statistical methods were developed to analyze the correlated data over 20 years. In this paper, we introduce a new class of parametric parsimonious distributions obtained from mixing negative binomial distribution with completely monotonic functions. By allowing the parameters to depend on covariates, we give a regression procedure that can be used to model the correlated data. We discuss maximum likelihood estimation and give asymptotic normality. The proposed procedure is applied to real data sets. Comparison is made with Poisson distribution model, generalized estimating equation model, and Williams' logistic linear model.

Maximum Likelihood Estimation of the Difference between Correlated Poisson Cough Counts

❖ Xin Yan, University of Missouri-Kansas City, 206 Haag Hall UMKC, 5100 Rockhill Road, Kansas City, MO 64110, *yanxi@umkc.edu*; Thomas Bradstreet, Merck & Co., Inc.

Key Words: Left truncated Poisson distribution, MLE, Correlated Poisson data

A statistical method is proposed to assess cough-modifying agents using the difference between cough counts before and after treatment. This clinical endpoint provides a direct assessment of efficacy in cough challenge studies. The conventional clinical endpoint is indirect, specifically the change from baseline in the concentration (either C2 or C5) of inhaled food-derived extract that triggers a predetermined number (either 2 or 5) of consecutive coughs in a given subject. Maximum likelihood estimation of the difference between cough counts before and after treatment is proposed assuming a left truncated correlated Poisson distribution. Estimation of the difference using simulated data is presented to illustrate the method. Two S-Plus functions are provided to facilitate calculating the maximum likelihood estimate of the difference and the corresponding confidence interval.

Negative Binomial Analysis of Hypoglycemia Rates in Diabetic Patients

❖ Cory Heilmann, Eli Lilly and Company, Lilly Corporate Center, DC 6036, Indianapolis, IN 46285, *heilmanncr@lilly.com*

Key Words: Negative binomial, Count data, Zero-inflation, Power, Poisson

Hypoglycemia rates in insulin using diabetic patients are difficult to quantitatively model due to the highly skewed nature of their distribution, as many episodes are experienced by a few patients while many patients do not experience any episodes. A common method of analyzing hypoglycemia rates is to model the rank transformed rates. However, a rank transformation dilutes the available information and complicates inference on the actual difference in rates. This study uses data from four clinical trials as well as simulation to test different models that analyze rates of hypoglycemia. Test statistics such as BIC from real data sets show superior fit of the negative binomial model to Poisson based models. Simulations from a variety of models show that negative binomial models have higher power than non-parametric models to detect differences without inflation of type I error rate.

Multivariate Dose-Response Analysis of Ordinal Response Data

❖Kao-Tai Tsai, Bristol-Myers Squibb Company, 34 Baldwin Dr, Berkeley Heights, NJ 07922, *tsai0123@yahoo.com*; Karl Peace, George Southern University

Key Words: Dose Response, Ordinal Response Data

In medical research clinical trials, patients are usually evaluated on multiple ordinal-scaled psychometric instruments. Conventional treatment efficacy assessment is conducted on each instrument individually. In this presentation, we consider a dose-response study in neuroscience and propose to assess the joint efficacy of the instruments based on a multivariate method which is applied to each stratum of the multivariate baseline measures. The performance of the proposed method will be compared with some conventional approaches. The significance of the proposed method will be evaluated using both the asymptotic distribution of the test statistic and the empirical permutation distribution of the data.

539 Multiplicity Issues and Possible Solutions in Clinical Trials ● ✪

Biopharmaceutical Section, Biometrics Section

Thursday, August 2, 10:30 am–12:20 pm

Gate-Keeping Testing Without Tears

❖Jianjun Li, Merck & Co., Inc., 351 N Sumenytown Pike, North Wales, 19454, *jianjun_li@merck.com*; Devan V. Mehrotra, Merck & Co., Inc.

Key Words: Gate-keeping procedures, trial-wise type-I error, multiple testing, primary endpoints, secondary endpoints

In a clinical trial, there are one or two primary endpoints, and a few secondary endpoints. When at least one primary endpoint achieves statistical significance, there is considerable interest in using results for the secondary endpoints to enhance characterization of the treatment effect. Because multiple endpoints are involved, regulators may require that the trial-wise type-I error rate be controlled at a pre-set low level. This requirement can be achieved by using “gate-keeping” methods. However, existing methods suffer from logical oddities such as allowing results for secondary endpoint(s) to impact the likelihood of success for the primary endpoint(s). We propose a novel and easy-to-implement gate-keeping procedure that is devoid of such deficiencies. Simulation results and real data examples are used to illustrate efficiency gains of our method relative to existing methods

Generalized Gate-Keeping Procedure

❖Ariful Haque, Novartis Pharmaceuticals, 529 Windmere Way, New Hope, PA 18938, *ariful.haque@novartis.com*; Sanat Sarkar, Temple University

Key Words: clinical trial, multiple tests, closed testing, gatekeeping strategy, Simes test

Two types of gate-keeping procedures are used in the analysis of clinical trails with multiple endpoints - serial and parallel. Serial gate keeping allows one to proceed to the testing of secondary family of hypotheses (related to secondary endpoints) only if all of the hypotheses in the primary or gatekeeper family (related to primary endpoints) are rejected. Whereas, parallel gate keeping allows one to proceed to the testing of secondary family of hypotheses when at least one of the hypothesis tests in the gatekeeper family exhibits significance. In this paper, we define a ‘generalized’ gate keeping strategy where one may proceed to the secondary family of hypotheses when at least k of the primary hypotheses in the gatekeeper family are found significant. We discuss the methods for constructing such a gate-keeping procedure using the generalized Simes test proposed by Sarkar (2005).

Multistage Gate-Keeping Procedures with Clinical Trial Applications

❖Alex Dmitrienko, Eli Lilly and Company, 10913 West 144th St, Overland Park, KS 66221, *alexei@lilly.com*; Ajit Tamhane, Northwestern University

Key Words: Multiple comparisons, Clinical trials, Type I error rate control

This talk introduces a general approach to constructing gate-keeping procedures for multiple testing problems arising in clinical trials with hierarchically ordered objectives (primary/secondary endpoints, dose-control comparisons, etc). The approach is applied to set up gate-keeping procedures based on popular multiple tests (Holm, fallback and Hochberg tests), resampling and parametric tests. The resulting procedures have a straightforward multi-stage structure that facilitates the implementation of gatekeeping procedures and communication of the results to non-statisticians. One can also account for logical restrictions among multiple analyses and improve the power of individual tests by eliminating comparisons that are no longer clinically meaningful. The general approach is illustrated using clinical trial examples.

Compatible Simultaneous Confidence Intervals for the Holm Procedure and Other Closed Bonferroni-Based Tests for One-Sided Multiple Comparison Problems

❖Frank Bretz, Novartis Pharmaceuticals, Basel, 4002 Switzerland, *frank.bretz@novartis.com*; Klaus Strassburger, Deutsches Diabetes-Zentrum DDZ

Key Words: multiple testing, multiple comparison procedures, confidence intervals, gatekeeping procedures, closed test procedure, partitioning principle

We consider the problem of simultaneously testing multiple one-sided null hypotheses. It is known that stepwise test procedures (for example, the Holm procedure) are by construction more powerful than their single step counterparts (for example, the Bonferroni test). This power advantage, however, comes only at the cost of increased difficulties in constructing compatible simultaneous confidence intervals for the parameter of interest. They are easily obtained for the Bonferroni method, but surprisingly hard to derive for the Holm procedure, for example. We discuss the inherent problems and show that ad hoc solutions used in practice typically do not control the pre-specified simultaneous confidence level. Instead, we derive simultaneous confidence intervals being compatible with a certain class of closed test procedures using weighted Bonferroni tests for each intersection hypothesis.

A Mixture Model Approach for Estimating the Proportion of True Null Hypotheses and Adaptive Control of FDR

❖ Ajit Tamhane, Northwestern University, Department of Statistics, 2008 Sheridan Road, Evanston, IL 60208, *ajit@iems.northwestern.edu*; Jiaxiao Shi, Northwestern University

Key Words: Beta Model, EM Algorithm, False Discovery Rate, Mixture Model, Normal Model, Adaptive Control of FDR

We study two mixture distribution models for the distribution of the test statistics with the goal being the estimation of the proportion coming from the null hypothesis and the use of this estimate in adaptively controlling the FDR. The normal model assumes that the test statistics follow a mixture of $N(0,1)$ and $N(\delta,1)$ with π_0 and $1-\pi_0$ as the mixing proportions. The beta model assumes that the p-values follow a mixture of $U[0,1]$ and Beta(a,b). Three methods of estimation are developed for each model. The methods are compared with each other and with Storey's estimator via simulation. Robustness is studied by generating data from other models. The EM algorithm performs best overall when the assumed model holds, but is not very robust to significant model violations. An example is given to illustrate the methods.

Evaluating Subgroup Risks in Trials with Many Subgroups: Issues and Potential Solutions

❖ Xiaojiang Zhan, Merck & Co., Inc., 126 E Lincoln Ave, Rahway, NJ 07065, *xiaojiang_zhan@merck.com*; Amarjot Kaur, Merck & Co., Inc.

Key Words: multiplicity, subgroup analysis, bias-correction, hazard ratio, estimation, clinical trials

In the analysis of clinical trials it is not uncommon to examine response to therapy in various patient subgroups in order to understand any differential in treatment response. ICH guidance for the common technical documents and many individual therapeutic area regulatory guidance also ask for subgroup analyses. In a typical subgroup analysis of risk, the estimates of risk are provided for multiple subgroups without accounting for multiplicity. Although the interest in examining multiple subgroups is valid to advance medical research, there is inherent cost (increase in Type I or Type II errors) associated with examining multiple subgroups. In this investigation the multiplicity issue is examined using time-to-event data. We characterize the magnitude of problem in terms of observed bias in estimating the hazard ratios for multiple subgroups and propose a potential correction.

Statistical Challenges in Proof-of-Concept Studies in Photosensitive Epileptics

❖ Norman Bohidar, Johnson & Johnson PRD, 445 Belle Lane, Harleysville, PA 19438, *norman.bohidar@comcast.net*; Geoff Edwards, Johnson & Johnson PRD; Liwen Xi, Johnson & Johnson PRD; Surya Mohanty, Johnson & Johnson PRD

Key Words: Epilepsy, Proof-of-Concept, Biomarkers

Photosensitive epileptics are commonly used as the target population for Proof-of-Concept studies for Epilepsy drugs. These patients are subjected to Intermittent Photic Stimulation (IPS), and EEG measurements are taken to define their photosensitivity range. Statistical challenges in these types of studies involve (1) deciding on decision rules for determining successful doses, (2) dealing with small sample sizes (3) handling multiple eye conditions and (4) using the responses on this biomarker in this model to predict successful doses in the general epileptic population. In the past, clinical judgment has been used to address these issues. This presentation will propose a statistical framework for evaluating them.

540 General Issues in Regression ●

ENAR

Thursday, August 2, 10:30 am–12:20 pm

The Theil-Sen Estimators in Modern Regression

❖ Hanxiang Peng, The University of Mississippi, PO Box 38677, University, MS 38677, *mmpeng@olemiss.edu*; Xin Dang, University of Mississippi

Key Words: depth function, efficiency, robustness, Theil-Sen estimator, semiparametric regression, Asymptotics

In this talk, we propose the use of multivariate medians to construct the Theil-Sen estimators for various regression models. We show they are robust, consistent and asymptotic normal. Simulations are conducted to compare robustness, efficiency and intensity.

Efficient Estimation of Population Quantiles in General Semiparametric Regression Models

❖ Arnab Maity, Texas A&M University, Department of Statistics, 3143, College Station, TX 77843, *amaity@stat.tamu.edu*

This paper considers a large family of semiparametric regression models, in which the main interest is on estimating population quantiles. Special cases of our framework include generalized partially linear models, generalized partially linear single index models, structural measurement error models and many others. We derive plug-in kernel-based estimators of the population quantiles of random variables and derive their asymptotic distribution. We also establish the semiparametric efficiency of these estimators under mild assumptions. We apply our methodology in an example in nutritional epidemiology where estimation of quantiles of usual intake of various dietary components are of primary interest. The generalization to the important case that responses are missing at random is also addressed.

Simulation Study Comparing Inverse Probability of Treatment Weighted (IPTW) Estimator and Multiple Regression Model

❖ Jianmin Wang, RTI Health Solutions, 200 Park Drive, Research Triangle Park, NC 27709-2194, *jmwang@rti.org*; William Irish, RTI Health Solutions

Key Words: Marginal Structure Model, Simulation, inverse probability of treatment weighted (IPTW) estimator, nonrandomized study, Multiple regression model

Bryan and colleagues (Biostatistics, 2004) compared naÔve and IPTW estimators of treatment effect using simulation when treatment assignment is related to confounders. The authors concluded that IPTW produces an unbiased estimate of treatment effect whereas a naÔve method is biased. Upon further review of the methods used, we found two serious limitations: 1) The naÔve estimator is based on a mis-specified model; and 2) Response variable is determined as a function of observed covariates, and therefore, has no random component. Using more appropriate methods, we conclude that the estimator based on a correctly specified model is asymptotically unbiased whereas the IPTW estimator is biased. This conclusion is in direct contrast to that reported by Bryan et al, and has important implications for the use of IPTW in the setting of nonrandomized studies.

Likelihood Ratio Test of Crossover Design Having Unequal Variances

❖ Yoon-Sung Jung, Kansas State University, 2142 Prairie Glen Pl, Manhattan, KS 66502, *ysjung72@gmail.com*; Dallas Johnson, Kansas State University

Key Words: likelihood ratio test, crossover design, unequal variance

We can consider a crossover design is a special form of repeated measures experiments. Appropriate structures of the variance-covariance structure need to have the correct analysis of variance. There is a problem of which parameter estimators does not have closed form solutions. An object in this paper is to show the likelihood ratio test of analysis of crossover designs when the variance-covariance structure has unequal variances.

Design Based on Interclass Correlation Coefficients

❖ Gheorghe Doros, Boston University, 715 Albany Street, Boston, MA 02118, *doros@bu.edu*

Key Words: Sample Size, ICC, Simulations

Reliability studies are concerned with the study of "consistency" or "repeatability" of measurements. The design of this type of studies is not well studied. For the case when adequate prior information is available, we propose a method for sample size calculation based on Interclass Correlation coefficients. The method assumes adequate prior information and uses simulations. A case study is presented.

A Locally Weighted Regression Approach to Evaluating the Effects of Smoking Reduction on Birth Weight

❖ Jeff M. Szychowski, The University of Alabama at Birmingham, RPHB 327, 1530 3rd Avenue South, Birmingham, AL 35294-0022, *jszychowski@ms.soph.uab.edu*; J. Michael Hardin, The University of Alabama; Michael D. Conerly, The University of Alabama; Lesa L. Woody, The University of Alabama at Birmingham; Wendy S. Horn, The University of Alabama at Birmingham

Key Words: smoothing, variable selection, smoking cessation

While smoking cessation is known to improve neonatal outcomes in pregnant smokers, reducing tobacco exposure has been more controversial. Windsor et al. (1999) showed that reducing tobacco exposure by 50% or more during pregnancy leads to beneficial effects on birth weight in neonatal outcomes. However, others have questioned the benefit of reduction in the absence of cessation. Some have argued that these conflicting reports are due to a nonlinear relationship between tobacco exposure and birth weight, (e.g., England, et al. (2001)). We employ a locally weighted regression (LWR) approach to evaluate the relationship between tobacco exposure and birth weight using data from the Smoking Cessation or Reduction in Pregnancy Trial (SCRIPT). By using an improved Cp statistic based on LWR, we verify this nonlinear relationship and identify a parsimonious predictive model for birth weight.

Modeling Lateral Position and Control in Driving Studies of Elderly Populations

❖ Jeffrey Dawson, The University of Iowa, Dept of Biostatistics C22 GH, 200 Hawkins Dr, Iowa City, IA 52242, *jeffrey-dawson@uiowa.edu*; Joseph Cavanaugh, The University of Iowa; K. D. Zamba, The University of Iowa; Matthew Rizzo, The University of Iowa

Key Words: Time series, Autocorrelation, Entropy

In studies of driver safety, high-frequency data collection allows the study of complex patterns of driving. For example, Boer (2000) and Dawson et al (2006) have used a measure called entropy to quantify lateral vehicular control. To investigate such approaches, specific models are needed to capture key aspects of driving data. We propose a model that accommodates the inertia of the vehicle path, correlated errors, and the environment wherein vehicles are used. We compare simulated data based on this model versus actual data from elderly subjects, showing which model settings best fit empirical data. We also describe random-effects extensions of this model to capture the between-person variability in driving performance. The use of these models may aid in assessing driver safety, and in the development of onboard systems to warn at-risk drivers of possible performance impairments.

541 Bayesian Nonparametric Modeling and Applications ●

Section on Bayesian Statistical Science

Thursday, August 2, 10:30 am–12:20 pm

Comparing Fits in Nonparametric Bayes Item Response Models

❖ Kristin Duncan, San Diego State University, 519 W Laurel St Unit 4, San Diego, CA 92101, *kristin.duncan@gmail.com*

Key Words: Nonparametric Bayes, Item Response Theory

Item response models are used to study the relationship between a latent trait or ability level and individuals' responses to items on a test or questionnaire designed to measure that trait. The item characteristic curve (ICC) gives the probability of a correct response as a function of ability. Dirichlet process priors provide a flexible means of modeling the ICC. Here we compare fits from models using two- and three-parameter logistic curves as the prior means for the Dirichlet process and also study the impact of the mass parameter of the Dirichlet process on the fitted curves. Comparisons are made with simulated data and with a data set from an introductory statistics final exam.

A Semiparametric Bayesian Approach to the Generalized Linear Mixed Effects Model

❖ Jing Wang, Louisiana State University, Baton Rouge, LA 70808, *jwang@lsu.edu*

Key Words: generalized linear mixed effects model, Dirichlet process, Gibbs sampler, Monte Carlo methods

We describe and illustrate Bayesian analysis in the generalized linear mixed effects model that allows the random effects to have a nonparametric prior distribution, using mixtures of Dirichlet process prior for the distribution of the random effects. The computations are implemented by using the Gibbs sampler; computational difficulties involved in numerical integrations in complex multiparameter structures are solved by Monte Carlo methods. The Examples using real data are given to illustrate the methodology.

Bayesian Inference for a Distribution-Valued Stochastic Process

❖ Zhen Wang, The Ohio State University, 2605 Lorain Ct, Columbus, OH 43210, *wangzhen@stat.osu.edu*; Steven N. MacEachern, The Ohio State University

Key Words: Dirichlet process

We propose two nonparametric Bayesian models that are natural extensions of traditional weighted least squares analysis. One model is appropriate for error distributions in a scale family; the other for data which are totals or averages. As the work extends weighted least squares, we consider a linear mean structure and take a model of normal errors as our starting point. The two models are indistinguishable in the parametric, normal-theory case. They become different when the error distributions are non-normal. For our models, the nonparametric component relies on a smoothed Dirichlet process prior with a normal base measure. Posterior inference is made on the basis of an efficient Gibbs sampler. For the first model, latent variables are introduced to facilitate computation.

A Bayesian Method for Curve Fitting

❖ Farideh Dehkordi-Vakil, The University of Iowa, N431F Dental Science Building, Iowa City, IA 52242, *farideh-dehkordi-vakil@uiowa.edu*; Nader Vakil, Western Illinois University

Key Words: Bayesian Estimation, Smoothing, Curve Fitting, Wiener Process

We propose a Bayesian nonparametric method for estimating smooth functions of bounded variation. Our method uses an integral representation of such functions and we use an integrated Wiener process to represent the prior information about the smoothness of the function. Using both simulated and real data, we illustrate our method and compare its performance with several competing methods. For our computations, we use the MCMC implemented in WinBUGS program.

Nonparametric Hierarchical Bayes Analysis of Binomial Data Using Berstein Polynomials

❖ Tingting Zhang, Harvard University, 1 Oxford Street 608, Cambridge, MA 02138, *tzhang@fas.harvard.edu*; Jun S. Liu, Harvard University

Key Words: Nonparametric Bayesian inference, Bernstein-DirichletBernstein-Dirichlet priors, Dirichlet process, Gibbs sampler

We propose a Bayesian nonparametric procedure for density estimation of binomial data with Bernstein-Dirichlet prior which is based on Bernstein polynomials. The predictive density is expressed as a weighted mixture of beta densities. This indicates that the predictive density is smooth everywhere, which is an improvement over the previous nonparametric Bayesian estimate with Dirichlet prior. The comparison between our approach and a previous nonparametric approach is provided through examples, and we find that Bernstein estimates are more robust to the sample variation than estimates with Dirichlet prior. Analysis of the bating data of Efron and Morris and the tack data of Beckett and Diaconis, which motivated this study, is supplemented to illustrate our method.

Distribution of Random Functional of a Ferguson-Dirichlet Process on the Unit Ball

❖ Thomas J. Jiang, National Chengchi University, Dept. of Math, 64 ChiNan Road Section 2, WenShan Taipei, 11605 Taiwan, *jiangt@nccu.edu.tw*; Kun-Lin Kuo, National Chengchi University

Key Words: Ferguson-Dirichlet process, c-characteristic function, spherical distribution, Fourier transformation

Jiang, Dickey, Kuo (2004) give the multivariate c-characteristic function and show that it has properties similar to those of the multivariate Fourier transformation. This new transformation can be useful when a distribution is difficult to deal with using Fourier transformation or traditional characteristic function. In this paper, we first give the multivariate c-characteristic function of the random functional of a Ferguson-Dirichlet process on the unit ball. We then find out its probability density function using properties of the multivariate c-characteristic function. This new result in three-dimension would generalize the two-dimensional result given by Jiang (1991).

542 Measurement Issues in Surveys ●

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Thursday, August 2, 10:30 am–12:20 pm

Comparison of the Income Items from the CPS and ACS

❖ Bruce Webster, Jr., U.S. Census Bureau, 9348 Cherry Hill Road, Apartment 104, College Park, MD 20740, *bruce.h.webster@census.gov*

Key Words: Income, Response Errors, Reference Periods, ACS, CPS ASEC

The Census Bureau conducts several surveys that collect income data. Each of these surveys gives slightly different results. This paper compares income responses given by the same respondent for the same time period on two separate surveys, the 2005 American Community Survey (ACS) and the 2006 Current Population Survey (CPS) Annual Social and Economic supplement (ASEC). To evaluate the sources of the differences in responses to the two surveys, the Census Bureau created a dataset that contains respondents who were in both surveys. This paper tries to isolate the reporting differences and provide possible reasons why those differences occurred.

The Impact of Converted Refusals upon Survey Estimates: National Health and Nutrition Examination Survey, 1999–2004

❖ Margaret Carroll, National Center for Health Statistics, CDC, 3311 Toldeo Road Room 4413, Hyattsville, MD 20782, *mdc3@cdc.gov*; Yinong Chong, National Center for Health Statistics

Key Words: Non-Response bias, Converted refusals

Converted Refusals have been hypothesized or proposed to decrease non-response bias of survey estimates. However, few studies have tested this hypothesis due to the lack of administrative data on these individuals. Beginning in 1999, the start of the continuous National Health and Nutrition Examination Survey, administrative data has been used to identify converted refusals. Focusing on adults 20 years and older we will address the issue of non-response bias by measuring the impact on converted refusals upon survey estimates. New sample weights will be constructed with the converted refusals treated as non-respondents. Based on these sample weights, the percent of adults with selected health characteristics will be estimated. These estimates will then be compared to corresponding estimates based on original sample weights where the converted refusals were treated as survey participants

Precanvass Survey for the 2007 Commodity Flow Survey

❖ Promod Chandhok, Bureau of Transportation Statistics, RITA, 8700 Ross Strret, Bowie, MD 20720, *promod.chandhok@dot.gov*

Key Words: auxiliary establishments, non-shipping auxiliaries, duplicates

The precanvass survey for the 2007 Commodity Flow Survey (CFS) will yield the single biggest improvement to the 2007 CFS. In the 2007 CFS, a sample of approximately 100,000 is selected from 750,000 establishments in select areas. Information will be collected from a sample of their shipments to produce estimates of the origin and destination of shipments, commodities shipped, their value, weight, and mode of transportation. The main accomplishments of the precanvass are 1) collected, for the first time in CFS history, measure of size information for auxiliary establishments;

2) eliminated more than 20,000 nonshipping auxiliaries; 3) eliminated more than 1000 duplicate addresses; 4) developed procedures for resolving unidentified as addressed; 5) updated contact information of precanvass establishments; and 6) collected information on the usage of third-party logistics providers.

An Appraisal of the Conduct and Provisional Figures of the Nigerian Population and Housing Census of 2006

❖ Raymond Okafor, University of Lagos, 169 Crownpoint Drive, Hydepark, MA 02136, *okaforray@yahoo.com*

Key Words: Nigerian Census of 2006, NPC, Conduct of census, Provisional figures, Controversy, Appraisal

This article appraises the conduct and provisional figures of the 2006 Nigerian population and housing census. A summary of the results, released in January, 2007, by the National Population Commission (NPC), is as follows: Nigerian population is about 140 million; populations of males and females are approximately 71.7 and 68.3 million, respectively and annual growth rate of about 3.2%. Most of the census figures are already being hotly contested and even rejected by many individual Nigerians, civil society groups and state governments. We dare say that controversies over census results are not new in Nigeria. Results from this research seem to indicate that current figures are following the same trend. Our findings seem to suggest that (a) a lot more could have been done during the conduct of the census, and (b) the NPC should appraise all complaints and take needed action(s).

Measuring the Quality of Service to Taxpayers in Volunteer Sites

❖ Kevin Cecco, Internal Revenue Service, 500 N Capitol St NW, Statistics of Income Division, Washington, DC 20001, *kevin.cecco@irs.gov*; Ronald Walsh, Internal Revenue Service; Rachael Hooker, Internal Revenue Service

Key Words: quality of service, Hansen-Hurwitz estimator, statistical consulting

There are nearly 12,000 locations nationwide where low-income, elderly, and military taxpayers can receive assistance in satisfying their tax responsibilities from non-Internal Revenue Service (IRS) volunteers. The Stakeholder Partnerships, Education and Communication (SPEC) organization of the IRS is responsible for building and maintaining partnerships with the stakeholders in local communities who oversee these locations. A key to maintaining these relationships is SPEC's ability to measure the overall quality of service provided by the volunteers who staff these locations. The Statistics of Income (SOI) Division of the IRS provides general statistical consulting services to various internal IRS customers. This paper will detail SOI's attempt to find a viable solution to SPEC's quality needs while balancing logistical issues, available resources, and customer expectations.

Statistical 'Pre- and Post-Processing' as Integrated into Computer Simulations

❖ Turkan Gardenier, Pragmatica Corp., 115 Saint Andrews Dr. NE, Vienna, VA 22180, *drgarden@verizon.net*

Key Words: Pre-processing, post-processing, simulation, Hadamard matrix, fractional factorial, response surface

Large scale computer simulations demand the generation of data in order to evaluate the effects of input parameters (singly and jointly) upon results. Design and selection of input data using Hadamard matrices and fractional factorials, also denoted as "pre- processing," enhance the reliability of outputs analyzed. Results may be expressed as a "meta-model" or submitted to stagewise refinement using statistical response surface methodology. In computerized simulation studies the process has sometimes been denoted as "post-processing." This presentation will illustrate the integration of a compendium of these statistical methods into large scale military simulations and their impact on cost reduction and data efficiency.

543 Nonparametric Inference for Functional Data

Section on Nonparametric Statistics, The ASA

Thursday, August 2, 10:30 am–12:20 pm

Functional Density Warping

❖ Zhen Zhang, University of California, Davis, 112 Koshland Way, Santa Cruz, CA 95064, *zhzhang@ucdavis.edu*; Hans-Georg Müller, University of California, Davis

Key Words: Age-at-death density, Hazard rate, Functional data analysis, Time warping, Quantile function, Kernel smoothing

When estimating an overall density function based on observations on individual densities for a sample of cohorts, varying scales between cohorts can be a problem which is ignored by conventional density estimation. The proposed density warping approach incorporates varying time scales between cohorts, yielding an overall density that more adequately represents the sample of cohort densities. Observed cohort densities are assumed to be time-warped realizations of an underlying stochastic process. In a first step, we obtain time-synchronized distribution functions from a warping mapping, from which both overall density and hazard functions are derived. Asymptotic properties of the resulting density estimates are derived, and we illustrate their performance in simulations and for longevity data obtained for 142 cohorts of Mexican fruit flies.

Asymptotic Theory of Bivariate Penalized Splines

❖ Yingxing Li, Cornell University, 301 Malott Hall, Ithaca, NY 14850, *yl377@cornell.edu*; David Ruppert, Cornell University

Key Words: Asymptotic bias, B-splines, Difference penalty, Number of knots, backfitting

We study the asymptotic behavior of penalized spline estimators in the bivariate case. Tensor-product B-splines are used and a penalty is placed on the differences of the coefficients. By choosing an appropriate penalty, the estimator has the same convergent rate as in the univariate nonparametric regression. Providing that the number of knots converges to infinity fast enough, the number of knots does not affect the asymptotic distribution. Expressions for the bias and variance are given. The results are compared to backfitting.

Local-Moment Polynomial-Smoothed Histograms with Application to Massive and Pre-Binned Datasets

❖ Galen Papkov, Rice University, 2300 Old Spanish Trl 2025, Houston, TX 77054, *gpapkov@rice.edu*; David W. Scott, Rice University

Key Words: Polynomial histogram, Nonparametric density estimation, smoothing, massive data sets, binned data, massive data sets, binned data

This paper suggests some logical and preferential modifications to Scott and Sagae's (1997) local moment method, the smoothed polynomial histogram (SPH), which uses a basis of truncated power functions (TPF) along with a roughness penalty to conduct nonparametric density estimation. Alternative tools presented in this paper include the addition of a weight matrix to be more faithful to information obtained from high-density bins, the use of B-splines which are more numerically stable than TPF, implementation of P-splines with a difference penalty which is simpler to code compared to a roughness penalty and increases computational efficiency, and quadratic programming to compensate for occasional negativ-

ity within the estimated densities. Finally, a procedure for extrapolating a smoothed polynomial histogram to higher dimensions via tensor products of B-splines is introduced.

Bayesian Isotonic Density Regression

❖Lianming Wang, National Institute of Environmental Health Sciences, 111 T. W. Alexander Drive, Biostatistics Branch, MD A3-03, Research Triangle Park, NC 27709, *wangl3@niehs.nih.gov*; David B. Dunson, National Institute of Environmental Health Sciences

Key Words: Conditional density estimation, Dependent Dirichlet process, Hypothesis test, Nonparametric Bayes, Quantile regression, Stochastic ordering

In estimating and performing inferences on conditional response distributions given predictors, stochastic ordering constraints can be incorporated to express prior knowledge and improve efficiency. This article proposes a nonparametric Bayes approach for modeling of an uncountable collection of stochastically ordered distributions indexed by a continuous predictor. Theory is developed to allow priors to be chosen with large support through a dependent Dirichlet process (DDP) specification. Choosing monotone splines for functional atoms within the DDP representation, an efficient MCMC algorithm is developed for posterior computation. Methods are developed for hypothesis testing and graphical presentation of results. The approach is evaluated through a simulation study and is applied to an epidemiologic study.

Spline Confidence Band with Mixing Error in Fixed Design

❖Jing Wang, University of Illinois at Chicago, 209 S Oak Park Ave, Apt 405, Oak Park, IL 60302, *wangjing@math.uic.edu*

Key Words: Polynomial Spline, Confidence Band, Mixing, Fixed Design

Base on polynomial spline regression, the asymptotically conservative confidence bands of the strong mixing time series are proposed in the cases of fixed design (equally spaced design). The explicit form is provided in the paper. Simulation experiments provide strong evidence that collaborates with the asymptotic theory. The result is applied to the Leaf Area Index data in East Africa to test the trigonometric curve assumption on the Leaf Area Index data which is collected regularly over years to improve the Regional Climate Model System.

Some New Normality Tests Using Kernel Density Estimators

❖Yishi Wang, Western Carolina University, Math and Computer Science Dept, Cullowhee, NC 28723, *ywang@wcu.edu*

Key Words: normality test, density estimators of convolutions

In this paper, we propose some new tests of normality using root-n consistent density estimators. Simulation shows that these tests has a better power compared with the test proposed in Arcones and Wang (2006). The consistency and asymptotic properties of the proposed tests will be discussed.

544 Data Analysis and Variable Selection ●

Section on Physical and Engineering Sciences, Section on Statistics and Marketing, Section on Teaching Statistics in the Health Sciences, Section on Quality and Productivity

Thursday, August 2, 10:30 am–12:20 pm

Variable Selection for LDA When n < p

❖Matthew Mitchell, Metabolon, PO Box 110407, Research Triangle Park, NC 27709, *mmitchell@metabolon.com*

Key Words: LDA, variable selection

Linear Discriminant Analysis (LDA) is a classification technique used for multivariate data. Each discriminant is a linear combination of all the variables. However, when the number of variables (p) is larger than the sample size (n), all variables cannot be used because the matrices in the discriminant function become singular. There are several methods for overcoming this difficulty: performing a principal component analysis (PCA), and then performing the LDA on the principal components; using univariate significance tests to choose a subset of the variables for the LDA; forward, backward, or stepwise variable selection; or more complex techniques. We compare these techniques in terms of their predictive ability and their ability to select the important features. These comparisons will be performed over several simulated data sets of various complexities and a real dataset.

Detecting Multiple Outliers and/or Mean Vector Shifts in Large Samples of Individual Multivariate Observations

❖Joe Sullivan, Mississippi State University, MS 9582, Starkville, MS 39762, *ymail397-asa@yahoo.com*

Key Words: Change point, minimum volume ellipsoid (MVE), retrospective analysis, robust estimation, statistical process control (SPC)

In a retrospective analysis of a historical data set, the data may have multiple sustained shifts and/or outliers. Especially with individual observations, subgroups of only one observation, the presence of out-of-control observations may disrupt parameter estimation, in turn masking their presence. The goal is partitioning the data so that all pairs of successive parts have statistically significantly different mean vectors and no finer partition has this property. A part is a set of observation(s) between consecutive boundaries. Robust estimators, such as the minimum volume ellipsoid, are computationally intensive for large amounts of data. Making use of the sequential nature of sustained shifts can reduce the computational burden. This research generalizes earlier univariate methods, presenting alternative approaches and preliminary comparison of their effectiveness.

Bayesian Wavelet-Based Despeckling of Ultrasound Medical Images Using the Gauss-Hermite Expansion

❖S. M. Mahbubur Rahman, Concordia University, Dept. of ECE, 1455 de Maisonneuve Blvd. W., Montreal, QC H3G 1M8 Canada, *mahb_rah@ece.concordia.ca*; M. Omair Ahmad, Concordia University; M. N. S. Swamy, Concordia University

Key Words: Ultrasound image, Despeckling, Wavelet coefficients, Gauss-Hermite, Prior function, Estimation

Wavelet-based despeckling of medical ultrasound images improves the diagnosis. Despeckling can be formulated as a statistical estimation technique, wherein the choice of an appropriate prior function of the sub-band wavelet coefficients is of major issue. In this paper, it is shown that a prior function that uses the Gauss-Hermite expansion performs better than standard ones, such as the generalized Gaussian or Bessel K-form, specifically because the proposed one can make use of an arbitrary number of higher order moments in parameter estimation. A Bayesian despeckling technique is then proposed in a homomorphic framework using the new prior function. Experiments are carried out using synthetically-speckled and real ultrasound images, and the results show that the proposed method performs better than several existing methods in terms of both the signal-to-noise ratio and visual quality.

Using an Adaptive Control Filter To Predict the Synchronization of Time Series

❖Morris Morgan, Hampton University, Olin Engineering Building, Hampton, VA 23668, *morris.morgan@hamptonu.edu*; Carolyn Morgan, Hampton University

Key Words: time series, adaptive control, chaos, stochastic time series

Past research used statistical measures (e.g., m-statistic, Kendall 'tau' statistic and permutation entropies) to assess the degree of synchronization in time series where the underlying dynamic models were known a priori. Order-based statistics that are robust to additive noise were used to control nonlinear oscillators, couple delayed equations and analyze nonlinear time series such as electrocardiogram data. Graphical techniques that quickly characterize such systems were investigated. Current work is directed at using an adaptive control procedure that provides statistical-based parameter estimates. The approach is applicable for series where the underlying models are not known but must be coupled dynamically in the presence of stochastic noise. An array of traditional filters (e.g., Butterworth, Chebyshev and Haar wavelet) will be used to improve the efficiency of this approach.

Statistical Aspects of Marketing Mix Models

❖Igor Mandel, Analytic Partners, 1342 Lexington Avenue, New York, NY 10128, *igor.mandel@analyticpartners.com*

Key Words: marketing mix models, regression analysis, factors selection, Monte Carlo simulation

The main tool used in marketing mix models is regression analysis. In this specific environment "over-studied" in general regression obtains several features, which are not typically considered in statistical literature in such a combination. These are direct character of influence of some variables to outcome and indirect of others; importance of time decaying effects; high correlation between factors, which should be used together because of economical requirements; desirable maximization of number of factors in a model, as opposite to traditional reductionism; a very big number of possible factors combined with small number of available data points; as a result of previous points, serious difficulties in optimal variables selection. Paper provides some recommendations on these issues with results of Monte-Carlo simulation about validity of regression coefficients.

Evaluation and Selection of Models for Out-of-Sample Prediction When the Sample Size Is Small Relative to the Complexity of the Data-Generating Process

❖Hannes Leeb, Yale University, Department of Statistics, 24 Hillhouse Avenue, New Haven, CT 06511, *hannes.leeb@yale.edu*

Key Words: model selection, out-of-sample prediction, generalized cross validation, S_p criterion, non-parametric regression, large number of parameters and small sample size

We study the problem of selecting a model that performs well for out-of-sample prediction. We do not assume that any of the candidate models under consideration is correct. Our analysis is based on explicit finite-sample results. The results are non-standard because we consider a situation where the sample size is small relative to the complexity of the data-generating process. Also, we allow for the case where the number of candidate models is (much) larger than sample size. For Gaussian data, we show under minimal assumptions that model selection based on generalized cross validation or the S_p criterion performs well in such situations, uniformly over large regions in parameter space. We also show that the performance of other model selectors, including AIC and BIC, can be anything from satisfactory or mildly suboptimal to completely unreasonable, depending on unknown parameters.

545 Regression ●

Section on Statistical Computing, Section on Statistical Graphics, IMS

Thursday, August 2, 10:30 am–12:20 pm

Simple Techniques for Estimation of Additive Rate Regression Models

❖M. Brent McHenry, Bristol-Myers Squibb Company, F209A 5 Research Parkway, Wallingford, CT 06492, *brent.mchenry@bms.com*; Stuart R. Lipsitz, Brigham and Women's Hospital; Debajyoti Sinha, Medical University of South

Key Words: exponential / piecewise exponential distributions, method-of moments, Poisson regression, additive hazard rate regression model

If the distribution of the failure time is exponential, Aitkin et al. (2005) showed that the MLE of the additive hazard regression model can be obtained using a Poisson linear model; we extend this result to the piecewise exponential distribution. However, when the hazard rates are close to 0 and/or censoring is high, there often exists convergence problems in the Newton-Raphson algorithm since the MLE is on or close to the boundary of the parameter space. Thus, we also propose a weighted least squares method-of-moments technique to obtain consistent and asymptotically normal regression parameter estimates. Simulations show that the weighted least squares estimate does not suffer from the convergence problems of the MLE. The piecewise exponential version provides estimates similar to semiparametric models, but are much simpler to compute than methods of Aalen (1989, 1993).

Alternative Penalty Functions for Principal Components

❖Trevor Park, University of Florida, 103 Griffin/Floyd Hall, P.O. Box 118545, Gainesville, FL 32611, *tpark@stat.ufl.edu*

Key Words: penalized likelihood, functional data, rotation

The high variability inherent in principal component analysis, particularly as applied to functional data, has spurred development of much methodology for rotating or regularizing the components to improve interpretation. We consider methods for constructing and applying penalty functions designed for special types of data and embracing various notions of interpretability. Effectiveness of specially-designed penalty functions is illustrated in the context of penalized likelihood principal component rotation, and some algorithms are described. Examples of application to data sets, including functional data sets, are offered.

A New Robust Partial Least Squares Regression Method (RoPLS) and Its Robustness

❖Asuman Turkmen, Auburn University, Auburn University, 222 Parker Hall, Department of Mathematics and Statistics, Auburn, AL 36849, *turkmas@auburn.edu*; Nedret Billor, Auburn University

Key Words: Partial least squares, Outlier, Robustness, Multicollinearity

Partial Least Squares (PLS) regression is an alternative to ordinary least squares (OLS) regression for relating observed responses to a set of explanatory variables where the explanatory variables are highly collinear and where they outnumber the observations. Ordinary PLS regression is known to be very sensitive to outlying observations since it is based on maximizing the sample covariance matrix between the response and a set of explanatory variables. Therefore, in this study, a robust PLS method (RoPLS), which is resistant to masking and swamping problems, is proposed. We also explore the robustness properties of the proposed robust PLS method.

Real and simulated data sets are used to compare the performance of the RoPLS with the existing methods.

Variable Selection in Factor Analysis Regression and Its Application

❖ Zenaida F. Mateo, University of Manitoba, Dept of Statistics, 338 Machray Hall, Winnipeg, MB R3T 2N2 Canada, *zeny_mateo@umanitoba.ca*

Key Words: backward procedure, factor analysis regression, variable selection, ordinary least squares regression, observed and predicted values

Factor Analysis (FA) regression is a statistical method of linear prediction when the dependent variables are subject to measurement errors. In this study, a backward elimination procedure is proposed for the selection of variables in FA regression. The squared correlation coefficient r squared (B-hat) was used as a criterion to measure the adequacy of a regression model and for evaluation and comparison of subset regression models. Some numerical investigation will be presented to analyze and evaluate the performance of the proposed variable selection procedure.

Purposeful Selection of Variables in Logistic Regression: Macro and Simulation Results

❖ Zoran Bursac, University of Arkansas for Medical Sciences, 4301 W Markham, Slot 781, Little Rock, AR 72205, *zbursac@uams.edu*; C. Heath Gauss, University of Arkansas for Medical Sciences; D. Keith Williams, University of Arkansas for Medical Sciences; David Hosmer, University of Massachusetts Amherst

Key Words: logistic regression, SAS PROC LOGISTIC, variable selection algorithm, purposeful selection, confounding

The main problem in any model-building situation is to choose from a large set of covariates those that should be included in the “best” model. A decision to keep a variable in the model might be based on the clinical or statistical significance. There are several variable selection algorithms embedded in SAS PROC LOGISTIC. Those methods are mechanical and carry limitations. Hosmer and Lemeshow (2000) describe a purposeful selection of covariates algorithm within which an analyst makes a variable selection decision at each step of the modeling process. In this paper, we introduce a macro, %PurposefulSelection, which automates this process. We conduct a simulation study to compare the performance of this algorithm with three well-documented variable selection procedures in SAS PROC LOGISTIC: FORWARD, BACKWARD, and STEPWISE. Results and implications are discussed in more detail.

A Realization of the Functionalized General Trichotomy in Logistic Regression

❖ Ligong Chen, Center for Prostate Disease Research, 1530 East Jefferson St, Rockville, MD 20852, *lchen@cpdr.org*; Yongmei Chen, Center for Prostate Disease Research; Jennifer Cullen, Center for Prostate Disease Research

Key Words: Functionalized General Trichotomy, Piecewise Regression, Logistic Regression, Double Thresholds, Convergence Rate of Residuals, Continuity Test

Based on symmetrically or asymmetrically iterative segmentation, we introduced a functionalized general trichotomy (FGT) for regression analyses. Convergence rates (CR) of means of combined absolute residuals were used as weights per iteration to directly estimate weighted means and confidence intervals of two thresholds. An asymptotic Chi-square statistic was constructed to test differences between the full and segmented models, as well as a statistic to test continuity, and a statistic to evaluate goodness of fit. A quadratic curve of the CR was identified and confirms that the thresholds corresponding to the maximal CR are biased and not optimal estimates for a population. The Bootstrap was also deemed to be of limited use for this method. A study sample obtained from a large public use dataset is analyzed with logistic regression to illustrate the computation and methodology.

546 Statistical Graphics: Methods and Applications ●

Section on Statistical Graphics

Thursday, August 2, 10:30 am–12:20 pm

Visualizing Cluster-Compressed Multivariable and Multialtitude Atmospheric Data

❖ Daniel Carr, George Mason University, Dept of Statistics MS4A7, Fairfax, VA 22030, *dcarr@gmu.edu*; Amy Braverman, Jet Propulsion Laboratory

Key Words: geospatial, cluster quality, JPL, AIRS

This talk addresses the challenge of visualizing cluster-compressed atmospheric data from AIRS (Atmospheric Infra-Red Spectroscopy) in a geospatial context. Each cluster has 35 variables including temperature and water vapor at 11 altitudes and cloud fraction at 10 altitudes. One approach uses Java shareware called CCmaps. CCmaps conveys three variables for regions in a map using sliders, color and conditioning. Generating regions representing both altitude and membership in latitude and longitude grid cells brings all altitude data for three variables into play. A second approach uses C++/OpenGL software called Glisten to represent data using 3D rendered glyphs. This shows altitude using a z-axis. Widgets enable variable selection and transfer function control of glyph features such a color (via color ramps), size, transparency, and filtering. Live examples convey options and insights.

An Exploratory Stroll Along the Beach

❖ Charlotte Wickham, University of California, Berkeley, 2545 Hillegass Ave, Berkeley, CA 94704, *cwickham@gmail.com*

Key Words: webcam, functional pca, environment, climate change, time series

Webcams monitoring New Zealand beaches for environmental change produce a huge amount of data. A single image can tell you what the weather is like, how big the surf is and how many people are sunbathing, but what can you learn when you have a collection of images from every hour of every day? Each pixel tells a story over time. How do we tell which ones are telling the same story? And how do the pixels relate back to the beach? I present an exploratory investigation of this data using functional data analysis. I investigate how treating a time series of images as functional data can help reveal the interesting bits. In particular, functional principal components can uncover pixels exhibiting similar variation over the time period observed. Interestingly, the scores on the components split the image into physically interpretable areas.

Characterizing Multivariate Data with High-Resolution Human Faces

❖ Dean Nelson, University of Pittsburgh at Greensburg, 1052 Route 259, Ligonier, PA 15658, *den@pitt.edu*; Joe Szurek, University of Pittsburgh at Greensburg

Key Words: Multivariate data, High-resolution human faces, Chernoff’s faces, Principal components

Over 30 years ago, Chernoff used cartoon images of faces to depict multivariate data. The development of technologies to capture and manipulate high-resolution images makes it unnecessary to limit our characterization of human faces to cartoons. An analogue of Chernoff's method can be applied directly to high-resolution images of human faces. Because this is done using a high dimension 'face space', transformations based on principal components can be used to characterize multivariate data with human faces that facilitate interpretation of a variety of multivariate statistical analyses (e.g., discriminant analysis, cluster analysis, and multidimensional scaling). Additional uses of the technique will also be presented.

Graphs in Social Science Texts: We Can and Should Do Better

❖ Naomi Robbins, NBR-Graphs, 11 Christine Court, Wayne, NJ 07470, *naomi@nbr-graphs.com*; Joyce Robbins, Touro College

Key Words: statistical graphs, Graphs in textbooks, Examples bad graphs

Students learn from what they see in textbooks. This presentation critiques graphs found in a sampling of introductory college-level sociology textbooks. These examples include graphs with unnecessary dimensions, with uneven scales, lacking explanations of reference lines and other elements, and that do not make the data stand out. In many cases, students first encounter statistical methods in social science classes and therefore it is crucial that the graphs they read follow up-to-date guidelines of proper data presentation. We believe that basic graphing principles should become part of social science curricula.

Longitudinal Multivariate Graphics in the Analysis of Time Management Data

❖ Jessica M. Scott, Brigham Young University, 726 E 1720 S, Orem, UT 84097, *jessiemae2000@aol.com*; Steven A. Wygant, Brigham Young University; Bruce Brown, Brigham Young University

Key Words: Time Series Scatterplot, Multivariate Graphs, Time Management

A dichotomization method is used to identify and contrast the time management practices of highly effective students, and those on academic probation. Male and female student comparisons are made, and also comparisons across majors. Hans Rosling's time series bivariate scatterplots are shown to be particularly effective in tracking and comparing patterns of time usage over the course of an academic year. The main limitation in Rosling's method is the restriction to two variables as they change over time. A method is presented for making time series scatterplots multivariate and displaying changes over time on three or more variables simultaneously. The method is illustrated in the comparisons of time usage patterns across the two academic effectiveness groups (effective and ineffective), for male and female students in each of the college majors.

Generating Data with Identical Statistics but Dissimilar Graphics: A Follow-Up to the Anscombe Dataset

❖ Sangit Chatterjee, Northeastern University, 16 Moraine Street, Jamaica Plain, Boston, MA 02130, *s.chatterjee@neu.edu*; Firat Aikut, Northeastern University

Key Words: Ortho-normalization, Genetic Algorithms, Non-linear Optimization

The Anscombe dataset (1973) is popular for teaching the importance of graphics in data analysis. It consists of four datasets that have identical summary statistics (e.g., mean, standard deviation, and correlation) but dissimilar data graphics (scatter plots). In this paper, we provide a general procedure to generate datasets with identical summary statistics but dissimilar graphics by using a genetic algorithm-based approach.

A Note on the Barnett-Cohen Censored Histogram

❖ Jong Kim, Portland State University, Dept of Math and Statistics, Portland, OR 97207, *jong@pdx.edu*; Bryan G. Schar, U.S. Census Bureau

Key Words: exploratory data analysis, left-truncated and right-censored data, nonparametric censored histogram, outlier detection, parametric censored histogram

Barnett and Cohen (2000, JCGS) extended the definition of the histogram to accommodate right-censored data. We elaborate their histogram. We first extend it to left-truncated and right-censored data, which often occur in medical studies. Next, we introduce a new parametric censored histogram based on log-location and scale distributions, which are commonly used for lifetime data modeling. Two measures of discrepancy between Kaplan-Meier estimator-based histogram and the proposed parametric censored histogram are automatically computed for each of the distributions to identify the true distribution in our simulation study. Results show the parametric censored histogram performs well. Once the true distribution is identified, this censored parametric histogram is also used to detect outliers. Real data examples are shown to illustrate the methods.

547 Estimation, Forecasting, and Hypothesis Testing ●

Social Statistics Section

Thursday, August 2, 10:30 am–12:20 pm

Examination of Two Issues Regarding Electronic Essay Scoring

❖ Sandip Sinharay, Educational Testing Service, Educational Testing Service, MS 12T Rosedale Road, Princeton, NJ 08541, *ssinharay@ets.org*; Shelby Haberman, Educational Testing Service; Jiahe Qian, Educational Testing Service

Key Words: Press statistics, linear regression, E-rater

Electronic essay-scoring usually involves the application of a linear regression model fitted on a sample of essays; the human rater score of an essay serves as the dependent variable and several numerical features of the essay serves as the independent variables in the regression. We examine two aspects of essay scoring. First, we focus on the problem of determination of the minimum sample size that will allow us to score essays with enough precision. This involves a study of the Press statistic that provides the error in predicting a new observation. Second, we examine when there is a need of human intervention in electronic essay scoring. For example, because of the use of a linear regression model, an ordinary essay with an extremely large value of a numerical feature may receive a high score. We employ outlier analysis to set up rules to flag essays with unusual numerical features.

Statistical Tests for Differential Functioning in Parametric Item Response Theory: A Monte Carlo Evaluation of Conventional DFIT Methods and a Bootstrap Alternative

❖ Gregory Petroski, University of Missouri-Columbia, School of Medicine, Office of Medical Research, Columbia, MO 65212, *petroskig@health.missouri.edu*; Steven J. Osterlind, University of Missouri-Columbia

Key Words: Item Response Theory, Test Bias, DIF, Differential Item Functioning

Differential test and item functioning (DIF/DTF) occur when individuals with the same ability but from different segments of the examinee population have different probabilities of success on a test or test item. Raju, et

al. (1992) proposed methods for detecting differential functioning of items and tests (DFIT) in the context of parametric item response theory for binary response items. The DFIT framework includes several tests of significance for DIF and DTF. A simulation study examined Type I error rates and power characteristics of the conventional DFIT tests and several bootstrap alternatives. Results indicate that the conventional tests have unacceptably high Type I error. A bootstrap test for DIF and DTF exhibited good control of Type I error and reasonable power when DIF items are excluded from the parameter equating step.

Forecasting Inflation: An Evaluation of the Performance of Approximate Factor Models

❖ Logan Kelly, University of Kansas, 2603 Belle Haven Dr., Lawrence, KS 66046, *lkelly@ku.edu*

Key Words: Factor Models, forecasting inflation, parameter instability

This paper evaluates the performance of approximate factor models in forecasting inflation in the United States by comparing simulated out of sample forecasts generated with approximate factor models to forecasts generated with both an auto-regressive process and a Philips curve. Approximate factor models are found out perform comparison models in terms of both root mean squared error and combined forecast regressions, but approximate factor models do not by themselves solve the problem of parameter instability.

Simulation Results of Bias Estimates for Compositional Data

❖ Don R. Warren, University of Texas at San Antonio, One UTSA Circle, Demography BB 4 06 30, San Antonio, TX 78249-0704, *don.warren@utsa.edu*

Key Words: Compositional Estimation

Compositional data are a partitioning subset of positive numbers within an data set observation that, because they sum to one, create difficulties in their estimation and interpretation. A substantial number of statisticians (particularly in applied areas) continue to ignore or reject the advice of specialists (a) to exercise caution in the interpretation of estimates based on compositional data and (b) to use special, minimally-biased methods of estimation. This article reports the results of simulations that address key issues of those who are skeptical of the need for the use of specialized compositional estimates.

Navigating the Multiplicity of Hypotheses in Evaluating Juror Selections

❖ Bruce Barrett, The University of Alabama, Box 870226, University Blvd., 300 Alston Hall, Tuscaloosa, AL 35487-0226, *bbarrett@cba.ua.edu*

Key Words: Jury selection, multiplicity, Batson challenge

The process of selecting a trial jury is usually accomplished with the alternate striking of prospective jurors by opposing counsel. In so doing, each side is prohibited from systematically excluding jurors on the basis of certain protected characteristics. The court decision, which established only race as a proscribed selection criterion, has since given rise to an expanding list that now includes gender, religion, national origin, and sexual preference. The trial judge must determine whether the observed strikes are inconsistent with a neutral consideration of these characteristics, and, finding so, take corrective action to ensure equal protection rights. We show that the multiple proscribed selection criteria create the statistical problem of multiplicity, and that the probability of a false signal under random selection quickly approaches 1.0 for common significance levels.

Hypothesis Testing with Lot Quality Assurance Sampling (LQAS) Using the Normal Approximation of the Binomial

❖ David Fitch, Universidad del Valle de Guatemala, Apartado Postal 82 01901, Guatemala, 01901 Guatemala, *dfitch@uvg.edu.gt*; Roberto Molina Cruz, Universidad Del Valle de Guatemala

Key Words: LQAS, Lemeshow, Valadez, USAID, Evaluation

Lemeshow and coauthors have papers on using probabilities from the binomial distribution to test a null hypothesis, e.g., that the child vaccination rate is not >80%. It would be more efficient to use normal approximations of binomial probabilities. With sample sizes of 28, probabilities for b(28, .8) are approximated using N[28x.8,28x.8(1-.8)] = N[22.4, 4.48]. The binomial probability of 27 or more vaccinated (1 or fewer unvaccinated) where the population rate is .8 or greater is .02 or less, hence Ho would be rejected using binomials, but with only 26 vaccinated it would not, the probability being .06. But using the normal, z=(26-22.4)/2.1166, = 1.70, and thus Ho would be rejected, the relevant probability being .0446. A concern is to see that correct statistics are being used to save children's lives. This is not the case with the Valadez LQAS method widely supported by USAID.

548 Statistical Methods and Applications in Epidemiology ●

Section on Statistics in Epidemiology

Thursday, August 2, 10:30 am–12:20 pm

Application of Johnson System of Distribution in Gene Selection

❖ Florence George, University of South Florida, 10716 Tavistock Dr, Tampa, FL 33626, *florajonet@yahoo.com*; Kandethody Ramachandran, University of South Florida

Key Words: differentially expressed genes, Johnson system of distributions

Microarrays have become increasingly common in biological and medical research. They enable the simultaneous study of thousands of genes and provide gene expression information on a whole genome level. A major goal of microarray experiments is to determine which genes are differentially expressed between samples. A new approach using the Johnson's system of distributions is proposed in this paper to make inference concerning which genes are differentially expressed, in which no specific parametric distribution is assumed for the gene expression levels. Here the null distribution of the summary statistic is estimated empirically using Johnson's system of distributions. The simulation results show that the new approach has a higher power over the typical traditional methods.

A Computational Approach to Variable Selection in Canonical Discriminant Analysis

❖ Man Jin, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, *mj2149@columbia.edu*

Key Words: Familial aggregation, syndromes, Rayleigh coefficients, epilepsy

In the genetic analysis of complex traits, it is often useful to define syndromes, constellations of traits that show high familial aggregation. The degree of familial aggregation for a linear combination of traits may be maximized by the Rayleigh coefficient. But with a large number of traits, using data to define a linear combination of traits may lead to over-fitting. Motivated by a family study of epilepsy conducted by Ottman (2006), we consider variable selection procedures to avoid over-fitting. Some approaches is proposed, and the approach is examined in simulation studies and in an application to the epilepsy family data.

Using Surrogate Outcomes for Improving Power To Detect Gene-Environment Interactions

❖ Tamanna Howlader, Concordia University, Dept. of Mathematics and Statistics, 1400 de Maisonneuve Blvd West, Library Building Room 901, Montreal, QC H3G 1M8 Canada, *tamanna@mathstat.concordia.ca*; Michal Abrahamowicz, McGill University; Yogendra P. Chaubey, Concordia University

Key Words: Gene-environment interaction, Surrogate outcome, Power

This study explores the use of quantitative surrogates of a clinical binary outcome to improve low power to detect gene-environment (G ◊ E) interaction which is of major concern in genetic-epidemiologic research. We consider hypothetical models of the relationship between the binary and quantitative surrogate outcomes, and their relationships to genetic susceptibility, exposure, and other risk factors. Simulations are used to estimate power of the test for G ◊ E interaction in linear and logistic regression models. Sensitivity analyses are performed to assess the impact on power of important parameters, such as strength of the underlying (G ◊ E) interaction effect, and measurement errors in the outcomes. It is found that under certain conditions, higher power can be achieved by replacing the binary outcome by a quantitative surrogate outcome.

Extension of the Cohort Analysis for Genetic Epidemiology Program To Assess Excess Risk of Cancer

❖ Mei Liu, The University of Texas M.D. Anderson Cancer Center, 1155 Pressler Street, Unit 1340, Houston, TX 77030, *meiliu@mdanderson.org*; Carol J. Etzel, The University of Texas M.D. Anderson Cancer Center; Michael Scheurer, The University of Texas M.D. Anderson Cancer Center; Melissa Bondy, The University of Texas M.D. Anderson Cancer Center; Sara Strom, The University of Texas M.D. Anderson Cancer Center; Christopher I. Amos, The University of Texas M.D. Anderson Cancer Center

Key Words: Glioma, Familial aggregation of cancer, Incidence, SEER, Standard Incidence Ratio

Cohort Analysis for Genetic Epidemiology (CAGE) is a program that calculates standardized cancer incidence ratios assuming the observed number of malignancies follow a Poisson distribution. The original database was the Connecticut Tumor Registry (CTR) (1935–1985), the oldest in the Surveillance, Epidemiology, and End Results (SEER) program. We updated the database to include CTR rates through 1995 and created a second database for all SEER data (1975–2000). The SEER Hispanic data was limited to 1991–2000, thus we imputed rates based on incidence rates for Whites. We applied our updated program to assess excess of cancer in relatives of 1648 glioma patients. We found significant excess of cancer among relatives of Whites and Blacks, but not for Hispanics, suggesting our imputation did not result in erroneous excess of cancer.

Length-Biased Sampling in Cancer Screening with Variable Test Sensitivity

❖ Sonya Heltshe, University of Colorado at Denver and Health Services Center, 800 W 12th Ave, Denver, CO 80204, *sonya.heltshe@uchsc.edu*; Karen Kafadar, University of Colorado at Denver and Health Services Center

Key Words: randomized screening trial, test sensitivity, length-bias, sojourn time, cancer

Length biased sampling exists in screening programs where disease may be detected during the preclinical stage, since a longer sojourn time yields a greater likelihood of being screen-detected (Zelen 1976). This paper quantifies the effect of length-biased sampling on clinical duration when cases are subject to periodic screening with variable test sensitivity. Screening test sensitivity as a function of age and sojourn time, underlying bivariate distribution, correlation (between sojourn and clinical time), mixing proportion of slow versus fast progressing disease, and the ratio of screening interval length to mean sojourn time all influence the magnitude of the effect that length-biased sampling has on the distribution of clinical duration.

Comparison of Adaptive Gaussian Quadrature, PQL, and Laplace 6 Estimates of Hierarchical Logistic Models When Comparing Small Incident Rates in Cluster-Randomized Trials

❖ Rafael Diaz, California State University, Sacramento, 2962 Claudia Ct, West Sacramento, CA 95691, *rdiaz@csus.edu*

Key Words: hierarchical logistic models, cluster-randomization with dichotomous outcomes, PQL, Laplace 6 approximation, Adaptive Gaussian Quadrature

Hierarchical logistic models allows for the comparisons of groups of proportions from cluster-randomized trials with binary outcomes. Parameter estimation techniques for these models can yield considerably different estimates when the cluster proportions are small. This study compares the commonly used Penalized quasi-likelihood technique, the relatively new Laplace 6 approximation, and the Adaptive Gaussian Quadrature technique used in mainstream statistical packages using Monte Carlo simulations. In these simulations the difference between two groups of small proportions, modeled after those commonly found in epidemiological interventions involving small incident rates, is compared by means of a hierarchical logistic model. Recommendations for the use of each of these techniques when comparing groups of small proportions are provided based on the results of this study.

Sequential McNemar Tests with Exact Methods

❖ Jiang Hu, Glotech Inc., 2559 Little Vista Ter, Olney, MD 20832, *hujia@mail.nih.gov*; Aiyi Liu, National Institutes of Health; Chengqing (Alan) Wu, National Institutes of Health

Key Words: McNemar tests, group sequential tests, two-stage designs, exact methods

We proposes exact group sequential methods to implement McNemar's test on matched pairs data, which are commonly encountered in medical studies. Both conditional and unconditional approach are studied and compared. The stopping boundaries, sample size and power of two stage tests are considered.

549 Statistical Methods in Health Science Research

Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences

Thursday, August 2, 10:30 am–12:20 pm

Overview of the Odds Ratio and Relative Risk: Inequality, Conversion, Efficiency, and Power of Testing

❖ Peng Liu, Food and Drug Administration, 5100 Paint Branch Pkwy, College Park, MD 20740, *peng.liu@fda.hhs.gov*; Debra A. Street, Food and Drug Administration

Key Words: Odds-risk inequality, Transformation, Relative Efficiency, Optimal Strategy

The odds ratio and relative risk are functionally related to the population incidence rate, exposure rate, the magnitude of risk, and the sampling design. In this presentation, we examine the statistical properties of the odds ratio and relative risk estimates as well as the variables affecting their inequality, efficiency and the power of testing when both estimators obtained from the same study and from two different studies. The implications of this study are: (1) the selection of the odds ratio or relative risk estimate should be based on which estimator possesses better statistical properties rather than on personal preference, and (2) the "efficiency of estimate" could be applied to devise the optimal sampling strategy, in particular, when the study involves a moderate or high incidence rate where the case-control approach may be no longer superior.

A Graphical Method for Assessing Mediating Variables Using Brunswik's Lens Model

❖ Matthew P. Spackman, Brigham Young University, 1001 SWKT, Provo, UT 84602, *matt_spackman@byu.edu*; Dawson Hedges, Brigham Young University; Bruce Brown, Brigham Young University

Key Words: multivariate graphics, lens model, Brunswik

The Brunswikian lens model (1956) is an effective tool for assessing the extent to which accuracy in a classification task is mediated by a set of measurements. The multivariate structure of the mediating variables in the lens model is substantially clarified by the use of multivariate graphics adapted to lens model analysis. The power of this multivariate graphical approach to the lens model is demonstrated with examples from research on the accuracy with which humans can identify emotions from voice samples, with acoustic variables as the mediators. Graphical lens model analysis is further demonstrated with a second example, using volumetric MRI data from various brain locations as mediators of the clinical accuracy of identification of post traumatic stress syndrome.

Determining Graft Survival After Clinical Islet Transplantation: Superior Performance of Extended Cox-Based Models

❖ Abdul Salam, University of Alberta, 8215 112 Street, 2000 CollegePplaza, Edmonton, AB t6g2c8 Canada, *abdul.salam@ualberta.ca*; Peter Senior, University of Alberta ; Abdulkadir Hussein, University of Windsor; Patricia M. Campbell, University of Alberta ; Kathleen LaBranche, University of Alberta; James A.M. Shapiro, University of Alberta

Key Words: survival analysis, extended Cox-based models, recurrent events, islet transplantation, panel reactive antibodies

Two or more islet transplants are often required to achieve insulin independence in patients with type 1 diabetes. Conventional outcome measures in clinical islet transplantation (CIT) are based on graft survival either from time of first or time of final transplant. Consequently, important data may be overlooked or imprecise estimates of graft survival result. Data from 91 patients undergoing 182 transplants were examined to determine the effect of pretransplant panel reactive antibodies on graft survival. Our objective is to illustrate, using real-life CIT data, the use of extended Cox-based models and compare with the Kaplan Meier method. We conclude that extended Cox-based models (such as AG; PWP-TT; PWP-GT; LWA) are potentially important in CIT data because estimates of hazard ratio (95%CI) are more consistent and precise compared with KM method.

BMI, Alcohol, and Colorectal Cancer

❖ Negasi Beyene, Centers for Disease Control and Prevention, POBox 1696, Hyattsville, MD 20788, *neb9@cdc.gov*

Key Words: Colorectal Cancer, BMI, Alcohol

Colorectal cancer is the third most common cancer and third leading cause of cancer-related mortality in the United States. Over the past decade, colorectal cancer incidence and mortality rates have modestly decreased. Until age 50, men and women have similar incidence and mortality rates; after age 50, men are more vulnerable. Different studies conclude differently if weight, height, and alcohol consumption are Colorectal cancer risk factors. Using the National Health Interview Survey (NHIS) 1997–2003 we will examine the rate of colorectal cancer and if BMI and alcohol have anything to do with having colorectal cancer or not. We also examine if there is difference incidence of colorectal cancer among different races/ethnicities.

Improving Inference on Diagnostic Tests Without Gold Standard When Auxiliary Information Is Available

❖ Gong Tang, University of Pittsburgh, 307 Parran Hall Biostatistics, 130 DeSoto Street, Pittsburgh, PA 15261, *got1@pitt.edu*

Key Words: Diagnostic test, Latent class model, Auxiliary information

Tumor grade is an important prognosticator in survival of cancer patients. Tumors are graded into three categories: well, moderate and poor under a complex system. However, the reproducibility is poor. With at least three readings, the prevalence and classification rates are estimable up to a permutation of these three categories under a latent class model (Dawid, 1979). Various structural models were imposed in the past to identify the categories when instrumental variables are available. Here a new method is proposed to identify the categories and classification rates when an auxiliary variable is known to be positively or negatively associated with the latent variable, e.g., the true tumor grade. This method is illustrated by analysis of tumor grade reading data from a joint study of the National Surgical Adjuvant Breast and Bowel Project (NSABP) and the Genomic Health Inc.

Comparisons of Statistical Methods in Analyzing Clinical Data Through Empirical Simulations

❖ Yongming Qu, Eli Lilly and Company, Lilly Corporation Center, Indianapolis, IN 46285, *qu_yongming@lilly.com*; Lei Xu, University of Wisconsin-Madison; Pandurang M. Kulkarni, Eli Lilly and Company

Key Words: Empirical simulation, resampling with replacement, statistical power, type I error

With increasing requirements, from both regulatory and scientific community, for prespecification of details of all analyses prior to unblinding of data in clinical trials, it is critical that one pre-specifies the most appropriate statistical method and model. Selecting a model based on assumption checking (after the data has been unblinded) either inflates type I error or compromises the statistical power to detect a difference. Previous research is mainly focused on comparing various analysis models either through simulation or case studies. We propose an empirical procedure that utilizes the historical data, evaluates various models of interest, and provides an optimal choice of models to be used for pre-specification in future studies. We applies this method to several commonly used laboratory variables and gives guidance about how to analyze these variables.

Identifying Patterns of Longitudinal Data Set with Meaningful Inflated Missing Values: A Case Study Combining Data Mining and Statistical Techniques

❖ Hua Fang, University of Nebraska-Lincoln, 539 N24th Street Apt 13, Lincoln, NE 68503, *jfang2@unl.edu*; Kimberly A. Espy, University of Nebraska-Lincoln; Maria L. Rizzo, Bowling Green State University; Honggang Wang, University of Nebraska-Lincoln

Key Words: inflated missing data, longitudinal study, two-part mixture model, clustering, data mining, growth pattern

Techniques for handling missing data exist separately in the fields of data mining and statistics. Methods for identifying patterns of inflated missing data in longitudinal studies are rare. In this research, an integrated approach is illustrated using a real observational data set where three types of missing data co-exist and account for a significant portion of the overall sample. Instead of using imputation methods, a two-part mixture model is introduced to model the inflated missing data and estimate the growth curves of each experimental subject over time. Based on individual growth parameter estimates and their auxiliary feature attributes, a clustering method is then integrated to identify the growth patterns. The combined approach exhibits the practical value of leveraging the statistical and data mining techniques in the current and future quantitative analyses.

Continuing Education

Courses

UTAH

Salt Lake City

July 28, 2007

CE_01C 8:30 am–5:00 pm

Bayesian Modeling in Practice

The ASA

Instructor(s): David Spiegelhalter, MRC Biostatistics Unit

There has been an extraordinary growth in the application of Bayesian methods in diverse fields - from fisheries to the regulation of medical devices, from mapping diseases to court cases. There are many articles, textbooks and some reasonable software, so that if you know what you want to do you can generally do it. However anyone wanting to conduct a Bayesian analysis is faced with choices for which rigid rules cannot be specified: for example, what type of model might be reasonable in this context, what kind of prior assumptions are reasonable, how to implement the analysis in software, how to monitor and check the computations, how to communicate the results, how to assess the importance and plausibility of all the assumptions and, most important, how to convince a skeptical audience that the analysis is reasonable? This course will deal with the practical aspects of Bayesian modelling, not only in terms of actually getting results out, and also how to deal with these difficult strategic questions. A wide range of applications will be used to illustrate the implementation of models of varying complexity, and we shall try and provide some tentative guidance on some of the more tricky aspects of Bayesian analysis. A particular aim is to show how we can acknowledge and explicitly include in a model realistic assumptions about potential inadequacies in the available data. The examples will be illustrated using the WinBUGS software. Prerequisite: We shall assume familiarity with basic probability theory and common distributions, as well as standard statistical concepts such as linear and non-linear regression, generalized linear models, random-effects modelling including generalized linear mixed models, meta-analysis, residual analysis, model comparison using deviance, and so on. A basic familiarity with Bayesian methods would be useful but not essential, and it would help considerably if you had installed WinBUGS and at least run through the tutorial. The most important prerequisite is an enthusiasm for carrying out statistical analysis that adapts to the evidence available, rather than forcing data into whatever standard packages can provide.

CE_02C 8:30 am–5:00 pm

Applied Longitudinal Analysis

The ASA

Instructor(s): Garrett Fitzmaurice, Harvard University

The goal of this course is to provide an introduction to statistical methods for analyzing longitudinal data. The main emphasis is on the practical rather than the theoretical aspects of longitudinal analysis. The course begins with a review of established methods for analyzing longitudinal data when the response of interest is continuous. A general introduction to linear mixed effects models for continuous responses is presented. When the response of interest is categorical (e.g., binary or count data), a number of extensions of generalized linear models to longitudinal data have been proposed. We present a broad overview of two main types of models: "marginal models" and "generalized linear mixed models." While both classes of models account for the within-subject correlation among the repeated measures, they differ in approach. Moreover, these two classes of models have regression coefficients with quite distinct interpretations and address somewhat different questions regarding longitudinal change in the response. In this course we highlight the main distinctions between these two types of models and discuss the types of scientific questions addressed by each. Prerequisite: Attendees should have a strong background in linear regression and some minimal exposure to generalized linear models (e.g., logistic regression).

CE_03C 8:30 am–5:00 pm

Design and Analysis of Crossover Experiments

Biopharmaceutical Section, The ASA

Instructor(s): Dallas Johnson, Kansas State University

Crossover designs are a special type of a repeated measures experiment where the experimental units are given different treatments in different sequences over time. They offer some real advantages over traditional designs when comparing different treatments, as crossover designs allow each experimental unit to serve as its own control. There are however certain pitfalls that must be avoided if crossover designs are going to be used effectively. This course is intended for those who are involved with analyzing certain kinds of messy experiments that involve crossover designs with ordinal and/or continuous data. The course will introduce the important issues associated with crossover experiments, and provide instructions as to how to use statistical packages such as SAS and SPSS to correctly analyze crossover experiments with and without carryover effects.

CE_04C 8:30 am–5:00 pm

Latent Class Analysis of Survey Error

Section on Survey Research Methods, The ASA

Instructor(s): Paul Biemer, RTI International/The University of North Carolina at Chapel Hill

A statistical framework for modeling and estimating classification error in surveys. It begins by examining some of the early models for survey measurement error (Census Bureau models; Kish model; etc.) and demonstrating their similarities, strengths and weaknesses. Then these models are cast in a general latent class modeling (LCM) framework where the true values of a variable are assumed to be unobserved (latent) and a survey response constitutes a single indicator of the latent variable. The parameters of the model include the target population proportions for a categorical variable to be estimated in the survey and the probabilities of misclassification; for example, for dichotomous variables, the false positive and false negative probabilities. Survey item reliability and construct validity as well as estimator bias are defined and interpreted in this context. Methods for estimating the model parameters and issues of model identifiability will be discussed. A number of examples and illustrations will be presented to demonstrate the estimation methods and the interpretation of the latent class analysis results. The utility of the models for evaluating and improving survey data quality will also be discuss and demonstrated. The course will introduce the students to the lEM software for fitting a wide-range of LCMs which can be downloaded from the Web at no charge.

CE_05C 8:30 am–5:00 pm

Statistical Methods for the Analysis of Recurrent Events

The ASA

Instructor(s): Richard J. Cook, University of Waterloo; Jerald F. Lawless, University of Waterloo

Recurrent event data arise in diverse fields such as public health, actuarial science, social sciences, economics, business and manufacturing. The aim of this workshop is to introduce a variety of models and statistical methods available for the analysis of recurrent event data. Models for the analysis of count data, gap time models, and intensity-based models will be discussed, as well as robust methods based on rate and mean functions. The emphasis will be on multiplicative models. Techniques for model assessment will be discussed. Features of the various models will be illustrated by application to several examples, taken from health research and reliability. Statistical analysis will be carried out using R/S-PLUS code. Prerequisite: Some familiarity with basic methods of survival analysis is required.

CE_06C 8:30 am–5:00 pm

Non-clinical Statistics for Drug Discovery

The ASA

Instructor(s): Kimberly Crimin, Wyeth; Thomas Vidmar, Pfizer Inc.

It is critical for the non-clinical statistician in the pharmaceutical industry to communicate effectively with scientists. They need to possess the ability to transform scientific questions into statistical hypotheses and models. The most appropriate statistical methods need to be utilized in order to answer those scientific questions. This means the tool bag of the non-clinical statistician needs to contain a broad range of statistical methods. This course will provide an opportunity for both new and experienced statisticians to enhance their statistical tools and improve their consulting skills. Since communication is the most important asset of a non-clinical statistician, the course begins by discussing the consulting aspect of the statistician in drug discovery. The class then proceeds to discuss the scientific questions, statistical hypotheses and the statistical methodologies used in the following areas: high-throughput screening, lead-finding animal experiments, drug safety, pharmaceutical sciences, chemometrics and genomics. In each area, the role of the consulting statistician will be emphasized. Prerequisite: The prerequisite knowledge for an attendee is a good understanding of linear models and multivariate methods. Definitions for scientific terms will be provided.

CE_07C 8:30 am–5:00 pm

Graphics of Large Datasets

Section on Statistical Graphics, The ASA

Instructor(s): Antony Unwin, University of Augsburg; Heike Hofmann, Iowa State University

Graphics are great for exploring data, but how can they be used for looking at the large datasets that are commonplace to-day? Large datasets bring new complications and require different emphases and approaches. This course discusses how to look at ways of visualizing large datasets, whether large in numbers of cases or large in numbers of variables or large in both. It is based on the new book *Graphics of Large Datasets*, Unwin AR, Theus, M., Hofmann, H. pub: Springer (2006). Data visualization is useful for data cleaning, exploring data, identifying trends and clusters, spotting local patterns, evaluating modeling output, and presenting results. It is essential for exploratory data analysis and data mining. Data analysts, statisticians, computer scientists, indeed anyone who has to explore a large dataset of their own, should benefit from attending this course. Recommended Textbook: Unwin A.R., Theus, M., Hofmann, H. Springer (2006). Graphics of Large Datasets. ISBN: 978-0-387-32906-2

July 29, 2007

CE_01C 8:30 am–5:00 pm

Bayesian Modeling in Practice

The ASA

Instructor(s): David Spiegelhalter, MRC Biostatistics Unit

There has been an extraordinary growth in the application of Bayesian methods in diverse fields—from fisheries to the regulation of medical devices, from mapping diseases to court cases. There are many articles, textbooks and some reasonable software, so that if you know what you want to do you can generally do it. However anyone wanting to conduct a Bayesian analysis is faced with choices for which rigid rules cannot be specified: for example, what type of model might be reasonable in this context, what kind of prior assumptions are reasonable, how to implement the analysis in software, how to monitor and check the computations, how to communicate the results, how to assess the importance and plausibility of all the assumptions and, most important, how to convince a skeptical audience that the analysis is reasonable? This course will deal with the practical aspects of Bayesian modelling, not only in terms of actually getting results out, and also how to deal with these difficult strategic questions. A wide range of applications will be used to illustrate the implementation of models of varying complexity, and we shall try and provide some tentative guidance on some of the more tricky aspects of Bayesian analysis. A particular aim is to show how we can acknowledge and explicitly include in a model realistic assumptions about potential inadequacies in the available data. The examples will be illustrated using the WinBUGS software. Prerequisite: We shall assume familiarity with basic probability theory and common distributions, as well as standard statistical concepts such as linear and non-linear regression, generalized linear models, random-effects modelling including generalized linear mixed models, meta-analysis, residual analysis, model comparison using deviance, and so on. A basic familiarity with Bayesian methods would be useful but not essential, and it would help considerably if you had installed WinBUGS and at least run through the tutorial. The most important prerequisite is an enthusiasm for carrying out statistical analysis that adapts to the evidence available, rather than forcing data into whatever standard packages can provide.

CE_08C 8:00 am–12:00 pm

Dose Finding in Drug Development

Biopharmaceutical Section, The ASA

Instructor(s): Naitee Ting, Pfizer Inc.; James MacDougall, Bristol-Myers Squibb Company

For dose finding study designs, the thinking process should start way before a drug candidate enters Phase II. Scientists need to make best use of data obtained from pre-clinical experiments and early Phase I trials, to learn from other drugs of the same class, and to understand the PK and PD properties of the drug candidate. It is also important to consider the formulation, the potential market environment and other similar drugs that may compete with the candidate under study. All these are important considerations for designing the dose response studies. After the dose response clinical data are collected, many statistical approaches are available to perform data analysis. The two main categories are modeling approaches and multiple comparison adjustments. Depending on stage of development and objectives of the study, appropriate method will be selected for data analysis. The usefulness of each approach under various situations will be discussed in this short course.

CE_09C 8:30 am–5:00 pm

Generalized Linear Mixed Models: Theory and Applications

The ASA

Instructor(s): Oliver Schabenberger, SAS Institute Inc.

The workshop discusses the theory of generalized linear mixed models and the application of these models. The course consists of two parts. Part I covers the requisite theory of generalized linear mixed models. Part II covers examples and applications and adds additional theoretical detail as needed; it is the majority of the course. The first part of the workshop makes the connection between linear models (LM), generalized linear models (GLM), linear mixed models (LMM), and generalized linear mixed models (GLMM) in terms of model formulation, distributional properties, and approaches to estimation. Participants learn that GLMMs are an encompassing family of models and come to understand the differences and similarities in approaches to estimation and inference within the class. The workshop discusses the pros and cons of various estimation approaches and describes their implementation with SAS/STATÆ software. The first part of the course ends with a discussion of over-arching issues the analyst must confront when working with correlated, non-normal data, such as managing overdispersion and using marginal versus conditional models. The second part of the workshop uses a variety of examples to revisit the theory taught in Part I, to develop new insights (e.g., low-rank mixed model smoothing), and to present applications from different disciplines. The applications range from modeling rates, proportions and counts with random effects, to GEE-type marginal models for non-normal data, to mixed model smoothing. A final section describes inferential procedures following parameter estimation, for example, estimation of complex linear hypotheses, multiplicity adjustments, and adjusted standard errors. Computations are based on the mixed model tools in SAS/STAT software, primarily the GLIMMIX procedure and the NLMIXED procedure.

CE_10C 8:30 am–5:00 pm

Statistical Evaluation of Medical Tests and Biomarkers for Classification

Section on Statistics in Epidemiology, The ASA

Instructor(s): Margaret Pepe, Fred Hutchinson Cancer Research Center; Holly Janes, Johns Hopkins University; Todd Alonzo, University of Southern California

Development of new biomarkers and medical diagnostic devices has rapidly accelerated. Their rigorous evaluation is a high priority for research. Yet principles and techniques for the design and analysis of these studies are not widely known. There are fundamental differences with methods for therapeutic and etiologic studies. Moreover, much basic methodology has developed recently. This course first covers estimation and comparison of Receiver Operating Characteristic (ROC) curves. We describe extensions to adjust for covariates that effect biomarker/test measurements. For assessing factors associated with test performance, ROC regression methods will be presented. We also consider how to evaluate the added benefit of a new test when standard tests or clinical variables already exist. Second, we consider the design of case-control studies that are most common in this field. Sample size calculations and optimal choice of case-control ratio will be presented. The attributes and limitations of matching controls to cases will be discussed. Third, prospective studies are considered. They are often subject to restrictions on ascertainment of the gold standard outcome variable. Finally, we discuss problems incurred when the gold standard reference test is itself subject to error. A new suite of freely available Stata programs will implement analyses. Prerequisite: Introductory statistics.

CE_11C 8:30 am–5:00 pm

Modeling and Data Analysis for Complex Surveys

Section on Statistics and the Environment, Section on Survey Research Methods, The ASA

Instructor(s): Jay Breidt, Colorado State University; Jean Opsomer, Iowa State University

Survey data often arise from the efforts of institutions (like government agencies) to describe characteristics of a heterogeneous population in a cost-effective way. This is a hard problem, with many operational constraints on the methods that can be used. Typically, institutions use complex surveys, with stratification, clustering, and unequal probabilities, and those surveys suffer from different rates of nonresponse among different subgroups. Researchers working with complex survey data may be tempted to use the variety of statistical methods at their disposal to answer research questions, but those methods often are not appropriate unless the complexities of the survey are explicitly taken into account. This short course is aimed at researchers with a basic background in statistical theory (elementary probability, sampling properties of estimators, likelihood) and methods (point and interval estimation, hypothesis testing, regression analysis) who need to analyze complex survey data. No previous background in survey sampling is assumed. We review the features that make survey data complex, including design properties and post-sampling adjustments. We describe and compare model-based and design-based approaches to estimation and inference with complex survey data, review available software, and illustrate with example data sets.

CE_12C 8:30 am–5:00 pm

A Practical Guide for the Design and Analysis of Virtual (Computer Model) Experiments

Section on Quality and Productivity, The ASA

Instructor(s): Jerome Sacks, National Institute of Statistical Sciences; William J. Welch, University of British Columbia

Computer models or "codes" are used throughout science and engineering to further understanding of fundamental processes, as replacements for time-consuming, costly, or infeasible physical experiments, and as tools for developing and designing products and policy decisions. A virtual experiment with a computer model involves making runs of the code according to some plan. Analysis of the output data from runs of the code requires statistical methodology that can address questions of (1) high-dimensionality (many model inputs), (2) small n (few runs), and (3) unknown non-linear complex relationships between inputs and outputs. In the last 20 years a systematic approach has emerged that treats these issues of experimental design and data analysis in the context of deterministic models (that is, computer models that produce the same output if the inputs are repeated). The purpose of the course is to describe these methods and provide a practical guide to their use, exemplified by explicit applications in engineering, science, and policy-making. Attendees should be acquainted with response surface methods and language as typified in the text "Introduction to Linear Regression Analysis" by D.C. Montgomery, E.A. Peck, and G.G. Vining.

CE_13C 8:30 am–5:00 pm

Data Quality and Record Linkage Techniques

The ASA

Instructor(s): Fritz Scheuren, NORC at the University of Chicago; William Winkler, U.S. Census Bureau; Thomas Herzog, U.S. Department of Housing and Urban Development

This course is based on the 2007 Springer book entitled Data Quality and Record Linkage Techniques. The course will provide a mixture of practical advice, mathematical rigor, management insight and philosophy. The intended audience is the working analyst, whether he or she is a survey statistician, census taken or business consultant. Our approach is to employ real life examples. Most illustrations come out of our successful practice. A few are contrived to make a point. Sometimes they come out of failed experience, ours and others. A main goal of the course is to aid practitioners to achieve a deeper understanding, at an applied level, of the issues involved in improving data quality through editing, imputation, and record linkage. The course begins with examples of quality successes and failures, motivating the techniques to be gone over in depth later. Basic editing techniques are discussed next, beginning with a short history but with more advanced editing and imputation techniques being covered as well. The heart of the course is our presentation of record linkage techniques. These, unlike most editing and imputation techniques, often can lead to real data quality improvements, rather than just mitigating the negative consequences of data errors. Course attendees are strongly encouraged to acquire and review textbook before attending course. Recommended Textbook: Herzog, Thomas N., Scheuren, F., Winkler, W. (2007) Data Quality and Record Linkage Techniques. Springer. ISBN: 978-0-387-69502-0

CE_14C 1:00 pm–5:00 pm

Adaptive Dose-Response Studies

The ASA

Instructor(s): Christopher S. Coffey, The University of Alabama at Birmingham; Brenda Gaydos, Eli Lilly and Company; José Pinheiro, Novartis Pharmaceuticals

Insufficient understanding of the dose response profile of a compound is a shortcoming of clinical drug development, often leading to incorrect dose selection for the confirmatory phase. Indeed, failure to characterize adequate dosing early is often cited as a key contributor to the high late-stage attrition rate currently faced by the pharmaceutical industry. Adaptive dose-response trials allow more efficient learning about the dose response, for both efficacy and safety, earlier in development, which should ultimately reduce overall costs/timelines and provide better dosing information. Such designs may be thought of as a subset of adaptive designs in general. The rapid proliferation of adaptive designs, and inconsistent use of terminology, has created confusion about the similarities and, more importantly, the differences among the techniques. This half-day course will clarify the differences between adaptive dose-response trials and other types of adaptive designs, review traditional fixed designs and adaptive dose-response designs, and provide information on developing a Bayesian adaptive dose-response trial. Real and simulated data examples will be used to illustrate the various methods discussed in the course. The course is intended for anyone with a basic understanding of statistical methods (at approximately the 2nd year of a graduate program), basic knowledge of Bayesian methods, and some exposure to drug development concepts.

July 30, 2007

CE_15C 8:00 am–12:00 pm

Categorical Data Analysis

The ASA

Instructor(s): Charles S. Davis, CSD Biostatistics

Categorical variables occur frequently in many types of research. The purpose of this course is to introduce, discuss, and provide examples of several approaches for analyzing data in which the response variable is categorical. The course will provide a reasonably comprehensive overview of the basic methods that are of most use to applied statisticians. It is not intended to be a state-of-the-art summary of research in this field. The topics to be discussed include methods for the analysis of two-way and stratified two-way contingency tables, logistic regression for binary and polytomous responses, weighted least squares analysis of categorical data, and loglinear models. Examples using SAS software will be presented. The prerequisite for this course is a general statistical education at the M.S. level (including a mathematical statistics course at the level of Hogg, Craig, and McKean, 2005, Introduction to Mathematical Statistics) and familiarity with regression and ANOVA models for continuous response variables).

CE_16C 8:30 am–5:00 pm

Multiple Comparisons and Multiple Tests

The ASA

Instructor(s): Peter Westfall, Texas Tech University

This short course will cover multiple comparisons and multiple testing methods, from basic and state-of-the practice to advanced and state-of-the-art. Topics include "family" choice, combined versus component tests, interval versus test-based inference, Intersection-Union versus Union-Intersection tests, False Discovery Rate versus Familywise Error Rate, and Bayes/frequentist views. Models include standard ANOVA, regression, and multivariate analysis, as well as more general models such as are used in survival analysis, GLIMs, mixed models, and nonparametrics. Current software implementations based on recent research in JASA by Westfall and Tobias will be emphasized. Applications include clinical trials, agriculture, business and genetics. Prerequisite: Knowledge of probability-based statistical inference through ANOVA, regression and applied multivariate analysis.

CE_17C 8:30 am–5:00 pm

Analysis of Clinical Trials: Theory and Applications

Biopharmaceutical Section, The ASA

Instructor(s): Christy Chuang-Stein, Pfizer Inc.; Alex Dmitrienko, Eli Lilly and Company; Geert Molenberghs, Hasselt University

The course covers four important topics that commonly face statisticians and research scientists conducting clinical research. The 4 topics are analysis of stratified data, multiple comparisons and multiple endpoints, interim analysis and interim data monitoring, and analysis of incomplete data. The course offers a well-balanced mix of theory and applications. It presents practical advice from experts and discusses regulatory considerations. The discussed statistical methods will be implemented using the SAS software. Clinical trial examples will be used to illustrate the statistical methods. The course is designed for statisticians working in the pharmaceutical or biotechnology industries as well as contract research organizations. It is equally beneficial to statisticians working in institutions that deliver health care and government branches that conduct health-care related research. The attendees are required to have basic knowledge of clinical trials. Familiarity with drug development is highly desirable, but not necessary. This course was taught at JSM 2005 and JSM 2006 and received the Excellence in Continuing Education Award in 2005. Recommended Textbook: Dmitrienko A, Molenberghs G, Chuang-Stein C, Offen W. (2005). Analysis of Clinical Trials Using SAS: A Practical Guide. SAS Publishing. ISBN: 978-1-59047-504-1

CE_18C 8:30 am–5:00 pm

Temporal Alerting Algorithms for Biosurveillance

Section on Statisticians in Defense and National Security, Section on Risk Analysis, Section on Quality and Productivity, The ASA

Instructor(s): Howard S. Burkom, Johns Hopkins University; David Banks, Duke University

This course will present concepts and application issues of statistical alerting algorithms for syndromic surveillance. Methods for prospective monitoring of both univariate and multivariate time series will be included. The public health context and selection of the input data streams will be discussed along with the consequent logistical constraints. The necessary background for statistical process control and data forecasting methodologies will be provided along with explanations of how these traditional approaches are adapted for routine public health implementation. Emphasis will be on applications, with discussion of how to manage seasonal and day-of-week effects and other temporal correlation issues, as well as customary data acquisition problems. Methods for evaluating algorithm detection performance will be presented with the goal of optimizing sensitivity to disease outbreaks at false alarm rates acceptable to public health investigation capability.

CE_19C 8:30 am–5:00 pm

Practical Bayesian Clinical Trial Design

Section on Bayesian Statistical Science, The ASA

Instructor(s): Peter F. Thall, The University of Texas M.D. Anderson Cancer Center

This short course will cover practical Bayesian methods for clinical trial design and conduct. Attendees should have at least a Masters degree in statistics, or equivalent experience, and an understanding of elementary Bayesian concepts. The course will include numerous illustrations using actual clinical trials. Most examples will be drawn from oncology, although the methods are broadly applicable. Methods for eliciting and calibrating priors, incorporating historical data, and using computer simulation to establish a design's frequentist properties will be illustrated. The morning will be cover phase I and phase I/II designs, including dose-finding using the continual reassessment method and logistic regression models, accommodating multiple ordinal toxicities weighted by their clinical importance, finding optimal dose pairs of two agents used in combination, dose-finding based on efficacy-toxicity trade-offs, optimizing schedule of administration rather than dose, and jointly optimizing dose and schedule. The afternoon will cover single-arm and randomized phase II trials, and hybrid designs that incorporate several methods. These will including a paradigm for monitoring multiple discrete outcomes, phase II designs based on event times, hierarchical Bayesian methods for trials with multiple disease subtypes, using regression to account for patient heterogeneity, adaptive randomization, and optimizing multi-stage dynamic treatment strategies.

CE_20C 8:30 am–5:00 pm

Classification and Regression Trees

The ASA

Instructor(s): Wei-Yin Loh, University of Wisconsin-Madison

In a classification or regression tree model, the data space is split into several partitions and a simple statistical model is fitted to each partition. Because the partitioning may be displayed as a tree structure, the model is intuitive and easy to comprehend. Further, because the partitions are highly adaptive, the best algorithms are capable of prediction accuracy at least as good as that of traditional methods. This course reviews the major algorithms, including C4.5, CART, CHAID, CRUISE, GUIDE, M5, and QUEST. Also covered are ensemble procedures, such as bagging, and various types of regression, including least squares, quantile, Poisson, logistic, and relative risk regression. The emphasis is on showing the strengths and weaknesses of the methods and their capabilities compared to non-tree methods. The methods are compared in terms of prediction accuracy, model interpretability, selection bias, and computational requirements. Examples are drawn from business, economics, medicine, engineering, science, sports, and other fields. Relevant software is discussed where appropriate. Attendees should be familiar with linear regression at the level of Weisberg's Applied Linear Regression and discriminant analysis at the level of Johnson and Wichern's Applied Multivariate Statistical Analysis.

CE_21C 1:00 pm–5:00 pm

Meta-Analysis: Statistical Methods for Combining the Results of Independent Studies

The ASA

Instructor(s): Ingram Olkin, Stanford University

Meta-analysis enables researchers to synthesize the results of a number of independent studies designed to determine the effect of an experimental protocol such as an intervention, so that the combined weight of evidence can be considered and applied. Increasingly meta-analysis is being used in the health sciences, education and economics to augment traditional methods of narrative research by systematically aggregating and quantifying research literature. A Google scholar search on meta-analysis plus different fields of research uncovered close to 200,000 hits in the social sciences (psychology, sociology, education), and a like number in medicine. Two meta-analytic examples are the effectiveness of mammography in the detection of breast cancer, and an evaluation of gender differences in mathematics education. The information explosion in almost every field coupled with the movement towards evidence based decision making, and cost-effective analysis has served as a catalyst for the development of procedures to synthesize the results of independent studies. In this workshop we provide an historical perspective of meta-analysis, discuss, some of the issues such as various types of bias and the effects of heterogeneity. The statistical methodology will include discussions of nonparametric and parametric models; effect sizes for proportions, fixed versus random effects, regression and anova models. New material on multivariate models will also be presented.

July 31, 2007

CE_22C 8:00 am–12:00 pm

Harnessing Data Streams through Statistical Computing

Section on Statistical Computing, The ASA

Instructor(s): Simon Urbanek, AT&T Labs - Research; Tamraparni Dasu, AT&T Labs–Research

Our tutorial is strongly motivated by JSM 2007's theme "Statistics: Harnessing the Power of Information". Data streams are a predominant form of information today, arising in areas and applications ranging from telecommunications, meteorology and rocketry, to the monitoring and support of e-commerce sites. Data streams are characterized by large volumes and high rates of accumulation. They pose unique analytical, statistical and computing challenges that are just beginning to be addressed. It is an important area that statisticians can make significant contributions to, an area rife with open research problems. In this tutorial, we give an introduction and overview of the analysis and monitoring of data streams. We discuss the analytical and computing challenges posed by the unique constraints associated with data streams. There are a wide variety of problems; data reduction, characterizing constantly changing distributions, detecting changes in these distributions, computing and updating models for evolving data streams, identifying outliers, tracking rare events, "correlating" multiple data streams and others.

The current work in this area is dominated by the computer science community, with a largely algorithmic approach with the emphasis on data queries. However it lacks analytical rigor and a strong theoretical framework, based mostly on disparate methodologies aimed at solving specific problems. Statisticians can make significant contributions in this area. Statistical computing is an ideal framework for the analysis of data streams. It offers statistical rigor and confidence guarantees, not just performance guarantees as is normally the case with algorithmic methods. We give an overview of existing literature and applications, highlighting opportunities for statistical research where appropriate. We make extensive use of examples and real life applications to elucidate the material. In particular, we discuss major applications that we have worked as running examples throughout the tutorial. We conclude with a discussion of open research problems in this dynamic area.

CE_23C 8:30 am–5:00 pm

Statistical Monitoring of Clinical Trials: A Unified Approach

Biopharmaceutical Section, The ASA

Instructor(s): Michael Proschan, National Institute of Allergy and Infectious Diseases

This course uses the B-value approach to show how to monitor many different types of trials, including those with continuous, dichotomous, or survival endpoints, with or without adjustment for covariates. This means that the same boundaries can be used for the different types of trials. Moreover, we show that calculations performed under the alternative hypothesis depend only on the expected value of the z-score at the end of the trial. We use this to show how to compute and interpret conditional and unconditional power, which are very useful in deciding whether to stop a trial for futility. We then cover different monitoring boundaries (including the flexible error spending function approach), inference following a group-sequential trial (p-values, estimation, confidence intervals), some Bayesian methods, and some adaptive sample size methods. Although the B-value paradigm is applicable in most situations, there are some in which it does not, such as when the sample sizes are too small. We discuss what to do in these cases as well. Recommended Textbook: Proschan, M.A., Lan, K.K.G., and Wittes, J.T. (2006). Statistical Monitoring of Clinical Trials: A Unified Approach. Springer. ISBN: 978-0-387-30059-7

CE_24C 8:30 am–5:00 pm

Dropout in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis

Biometrics Section, The ASA

Instructor(s): Michael Daniels, University of Florida; Joseph W. Hogan, Brown University

This course provides a survey of modern model-based approaches to handling dropout in longitudinal studies, and illustrates the use of newly-developed methods for sensitivity analysis and incorporation of prior information. The emphasis is on Bayesian approaches but the models and methods discussed can be implemented in non-Bayesian settings as well. The course will begin with a brief review of models for longitudinal data and the basics of Bayesian inference. Included in this will be a quick primer on the WinBUGS software, used throughout the course to illustrate the concepts and models on real data examples. The second part of the course will focus on dropout. We will discuss formal classifications of the dropout mechanism and describe different classes of models to adjust for biases caused by dropout. We also will discuss the importance of model specification and the logistics of model fitting. Models for both 'ignorable' and 'nonignorable' dropout will be covered. The final part of the course focuses on nonignorable dropout; we will describe and motivate principles that should guide assessment of sensitivity to missing data assumptions and appropriate use of prior information. Attendees should have working knowledge of generalized linear models and statistical inference at the masters level.

CE_25C 8:30 am–5:00 pm

Statistical Genetics

Biometrics Section, The ASA

Instructor(s): Kenneth Lange, University of California, Los Angeles; Janet Sinsheimer, University of California, Los Angeles; Eric Sobel, University of California, Los Angeles

This course will present an overview of statistical genetics with emphasis on methods for detection of genetic loci for complex traits. The course will be balanced between theory and applications. After a brief but thorough refresher on the biological principles underlying statistical genetics, we will cover statistical models that underlie many pedigree-based gene mapping methods. We will then cover the statistical theory and implementation of methods for: allele frequency estimation, estimation of ethnic admixture, transmission distortion and gene mapping, penetrance estimation using generalized linear models, variance component models for quantitative trait mapping, and variance component models for inbred strains. Prerequisite: Familiarity with basic statistical modeling and likelihood theory is assumed. Recommended Textbook: Lange, K (2002). Mathematical and Statistical Methods for Genetic Analysis, 2nd edition. Springer-Verlag. ISBN: 978-0-387-95389-2

CE_26C 8:30 am–5:00 pm

Wavelets and Statistical Applications

Section on Bayesian Statistical Science, The ASA

Instructor(s): Brani Vidakovic, Georgia Institute of Technology; Marina Vanucci, Texas A&M University

Statistical wavelet modeling and computational research has, in this recent decade, become a burgeoning area in both theoretical and applied statistics, and is, in many cases, guiding the development of statistical methodologies in various interdisciplinary fields. Wavelet ideas are developing in statistics in areas such as regression, density and function estimation, linear models, spatial statistics, functional data analysis, time series, and others, with a number of applications in science and engineering. Wavelets ``love'' massive data sets and are especially efficient in data reduction, assessment of scaling, signal separation and data whitening. They are an indispensable tool in a data-miner's toolbox. This course will focus on the use of wavelet methods in statistics. The course will start with a brief introduction to the wavelet theory, followed by an overview of wavelet-based statistical methods and applications. Topics will include smoothing of noisy signals, nonparametric estimation of functional data and representation of stochastic processes. Emphasis will be given to Bayesian inferential procedures. Matlab software will be used. The course is aimed at statisticians and scientific investigators who are interested in the practical use of wavelets for the analysis of data. No knowledge on wavelets is required. Complex mathematical details will be only partially covered. Recommended Textbook: Vidakovic, B (1999) Statistical Modeling of Wavelets. Wiley. ISBN: 978-0-471-29365-1

CE_27C 8:30 am–5:00 pm

Analysis of Censored Data

The ASA

Instructor(s): Danyu Lin, The University of North Carolina at Chapel Hill

Many scientific studies are concerned with the times to the occurrences of certain events. A common complication is that the event times are censored on some study subjects due to drop-outs or study termination. There have been tremendous developments of statistical theory and methods for analyzing censored event time data in the last three decades. The purpose of this short course is to review some of the key developments. We will first provide a heuristic introduction to the counting-process martingale theory. We will then use this powerful tool to understand the properties of commonly used statistics, such as Kaplan-Meier estimator, (weighted) log-rank tests and Cox proportional hazards regression. We will address practical issues in applying the Cox model to real studies, including variable selection, model checking, model misspecification, and robust inference. We will also review recent advances in the field, including analysis of multiple events data, informative drop-out, and joint modeling of repeated measures and event times. Relevant software will be described. A number of clinical and epidemiologic studies will be provided for illustrations. The materials will be presented at a non-technical level. Although cutting-edge research will be discussed, this course is targeted primarily at applied statisticians (particularly clinical trial statisticians) who wish to analyze their data with the best available methods. Basic knowledge in mathematical statistics and linear models is required. No prior knowledge of survival analysis is necessary. There is no required textbook. A useful reference is "The Statistical Analysis of Failure Time Data" by J. D. Kalbfleisch and R. L. Prentice (2002).

CE_28C 1:00 pm–5:00 pm

Bootstrap Methods and Permutation Tests for Doing and Teaching Statistics

Section on Statistical Education, Section on Statistical Computing, The ASA

Instructor(s): Tim Hesterberg, Insightful Corporation

Early in Stat 101 we teach that robustness is important. Yet later in the course, and too often in statistical practice, we ignore those lessons, and use simple means and least-squares regression together with Normal-based inferences, even though the corresponding assumptions are violated. Bootstrapping and permutation tests (BPT) let us check the accuracy of common procedures, and the results are surprising. We'll see how inac-

August 1, 2007

CE_29T 8:00 am–9:45 am

Introduction to CART: Data Mining with Decision Trees

ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This course, intended for the applied statistician wanting to understand and apply the CART methodology for tree-structured nonparametric data analysis, will emphasize on practical data analysis involving classification. All concepts will be illustrated using real-world examples. The course will begin with an intuitive introduction to tree-structured analysis. Working through examples, we will review how to read CART output and set up basic analysis. This session will include performance evaluation of CART trees and cover ways to search for possible improvements of the results. Once a basic working knowledge of CART has been mastered, we will focus on critical details essential for advanced CART applications, including choice of splitting criteria, choosing the best split, using prior probabilities to shape results, refining results with differential misclassification costs, the meaning of cross validation, tree growing, and tree pruning. The course will conclude with discussion of the comparative performance of CART versus other computer-intensive methods, such as artificial neural networks and statistician-generated parametric models.

CE_30T 8:00 am–9:45 am

New Features for Clustered and Survey Data Analysis in SUDAANÆ Release 9.0.3

ASA

Instructor(s): Angela Pitts, RTI International; G. Gordon Brown, RTI International

This workshop will highlight 10 new features in SUDAAN Release 9.0.3 available to researchers analyzing cluster-correlated or complex survey data. These features include CMH Test of Trends; addition of new test statistics, including the Rao-Scott test; small proportion confidence interval estimation; introduction of a weighted kappa; chi-squared test to known proportions; updated memory manager; revised PRINT capabilities; logrank test in KAPMEIER; percentile correction in PROC DESCRIPT; and INCLUDE option in every procedure. The workshop will include a brief review of statistical problems encountered in survey data, plus a brief introduction to SUDAAN syntax. Attendees are not required to be SUDAAN users, but should have a good knowledge of statistical issues present in correlated data and surveys and be familiar with commonly used statistical methods and categorical data analysis. The new SUDAAN features will be demonstrated on complex survey data. We will demonstrate proper implementation of SUDAAN, provide interpretation of the output, and discuss statistical issues related to the data. The datasets, SUDAAN code used in the presentation, and a 30-day trial version of SUDAAN Release 9.0.3 will be provided to attendees.

CE_31T 8:00 am–9:45 am

Modern Regression Analysis in SAS Software

ASA

Instructor(s): Robert Cohen, SAS Institute Inc.

This workshop is intended for a broad audience of statisticians and data analysts who are interested in modern regression methods. I will describe these methods and SAS tools for fitting robust local regression models with the LOESS procedure, robust parametric models with the ROBUSTREG procedure, quantile regression models with the new QUANTREG procedure, mixed model smoothing with the new GLIMMIX procedure, and LASSO/LAR regression models with the new GLMSELECT procedure. These tools are illustrated with several examples. Attendees should have a basic understanding of regression theory.

CE_32T 8:00 am–9:45 am

Power Analysis: A Simple and Effective Approach

ASA

Instructor(s): Hannah Rothstein, ; Michael Borenstein,

This workshop offers an introduction to power analysis and shows how to use it simply and effectively to plan studies and write grant applications. Power analysis is the art of finding and justifying a sample size for a planned study. Power is determined by such factors as the magnitude of the treatment effect, the sample size, alpha, and (for survival studies) the study duration. The goal of the power analysis is to balance these as a series of "what ifs," done most effectively with graphs that allow the researcher to grasp (and communicate) a range of options in a single picture, and then find the one that strikes the optimal balance of feasible sample size, reasonable assumptions, and acceptable power. We will show how to do this using Power And Precision. We will use examples from means, proportions, and survival analysis. This workshop is recommended for people new to power analysis, frustrated by power analysis, and who need to perform power analysis for survival studies.

CE_33T 8:00 am–9:45 am

EastÆ 5: A Comprehensive Package for Adaptive and Group Sequential Design, Interim Monitoring, and Simulation

ASA

Instructor(s): Cyrus Mehta, Cytel Inc.

We will release EastÆ 5, a major upgrade of the East software package for design, interim monitoring, and simulation of adaptive and group sequential clinical trials. New analytical and simulation capabilities are provided for the survival setting that can accommodate fixed and variable follow-up, nonuniform accrual, varying hazard rates, and drop-outs. The software will support both group sequential designs, where early stopping is the goal, and adaptive designs, where sample size re-estimation is the goal. For the sample size re-estimation problem, we will demonstrate how East can compute confidence intervals and point estimates following an adaptive design. Finally, we will demonstrate a new East module for adaptive dose-response modeling.

CE_34T 10:00 am–11:45 am

Introduction to MARS: Predictive Modeling with Nonlinear Automated Regression Tools

ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This workshop will introduce the main concepts behind Jerome Friedman's MARS, a modern regression tool that can help analysts quickly develop superior predictive models. MARS is a nonlinear automated regression tool that can trace complex patterns in the data. It automates the model specification search, including variable selection, variable transformation, interaction detection, missing value handling, and model validation. Conventional regression models typically fit straight lines to data. Although this usually oversimplifies the data structure, the approximation is sometimes good enough for practical purposes. However, in the frequent situations in which a straight line is inappropriate, an expert modeler must search tediously for transformations to find the right curve. MARS approaches model construction more flexibly, allowing for bends, thresholds, and other departures from straight lines from the beginning. Attendees will be presented with the key benefits over conventional regression tools and over a modelers' tedious search for transformations to find the right curve.

CE_35T 10:00 am–11:45 am

From Software to Solutions in Statistics and Risk Analysis

ASA

Instructor(s): Shawn Harahush, Palisade Corporation

The world of business and education has become more complex with the decision of what software a business will use to successfully manage their incoming data. @RISK integrates into Microsoft Excel to provide a powerful Monte Carlo simulation engine for the ease-of-use environment. NeuralTools adds sophisticated neural networks analysis to Excel, and StatTools also integrates with Excel.

CE_36T 10:00 am–11:45 am

Design and Analysis of Split-Plot Experiments Using the SAS ADX Interface

ASA

Instructor(s): Sharad Prabhu, SAS Institute Inc.; JosÈ G. RamÌrez, W.L. Gore & Associates, Inc.

The SAS ADX Interface is a point-and-click solution for statisticians, engineers, scientists, and other researchers who design, analyze, and interpret experiments to improve industrial processes and products. This workshop will introduce the ADX Interface, focusing on the design and analysis of split-plot experiments. We will review the basic concepts of split-plot designs and explain how to use the ADX Interface to construct and analyze various types of fractional-factorial split-plot designs. Examples will demonstrate the many interactive and graphical features of the ADX Interface, which requires no knowledge of SAS programming.

CE_37T 10:00 am–11:45 am

Meta-Analysis: Concepts and Applications
ASA

Instructor(s): Michael Borenstein, ; Hannah Rothstein,

We will explain the concept of meta-analysis and discuss common criticisms, such as "apples and oranges," and conflicts between meta-analyses and clinical trials. We'll explain how to compute treatment effects, synthesize data using various computational models, use ANOVA and meta-regression, create forest plots, and address the potential impact of publication bias. We will show how to accomplish these tasks using Comprehensive Meta Analysis Version 2 (Borenstein, Hedges, Higgins, Rothstein), a new program developed with funding from the National Institutes of Health. The course is intended for people who need to perform or interpret meta-analyses in their work. Attendees should have some familiarity with meta-analysis, but the course will cover the basics before moving to more advanced topics.

CE_38T 10:00 am–11:45 am

Generalized Linear Modeling in SPSS
ASA

Instructor(s): Weicai Zhong, SPSS Inc.

GENLIN covers a variety of exponential distributions to analyze different data-including normal, gamma, inverse Gaussian, Poisson, negative binomial, and binomial-and many types of link functions-including identity, complementary log-log, log, log complement, logit, negative binomial, negative log-log, odds power, probit, and power. In this workshop, we will introduce model details and show how to choose and fit a generalized linear model using examples of data in applied research. We will demonstrate assessing the fit of a model, comparing it against competing models, interpreting the parameters, and performing follow-up analysis, such as estimated marginal means or scoring. Examples will illustrate the use of GENLIN. Attendees are expected to have basic knowledge of statistical inference. Familiarity with classical linear modeling is desirable, but not required.

CE_39T 1:00 pm–2:45 pm

Advances in Data Mining: Jerome Friedman's TreeNet/MART and Leo Breiman's Random Forests
ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This workshop will present Leo Breiman's Random Forests and Jerome Friedman's TreeNetT/MART (also known as TreeNet Stochastic Gradient Boosting). Random Forests and MART/TreeNet are new advances to classification and regression tree software, which enable the modeler to construct predictive models of extraordinary accuracy. Random Forest is a tree-based procedure that makes use of bootstrapping and random feature generation. In TreeNet, classification and regression models are built gradually through a potentially large collection of small trees, each of which improves on its predecessors through an error-correcting strategy. I will show how the software is used to solve real-world data mining problems, cover theory and discuss what is novel in the software, cover implementation, compare the two methodologies, and show where the software fits in terms of other data mining software.

CE_40T 1:00 pm–2:45 pm

Quantile Regression Using the SAS QUANTREG Procedure
ASA

Instructor(s): Colin Chen, SAS Institute Inc.

This workshop will introduce the SAS QUANTREG procedure, which computes conditional quantile functions and conducts statistical inference on regression parameters. Participants will learn to explore the heterogeneity of both cross-sectional and longitudinal data and interpret quantile effects and quantile processes. Basic and advanced features of the procedure will be presented with examples, including the construction of growth charts for medical measurements. New features, including quantile regression splines, also will be demonstrated.

Continuing Education

Instructor Index

Salt Lake City

UTAH

INDEX OF INSTRUCTORS

Index
of
Participants

UTAH

Salt Lake City

July 29 – August 2, 2007

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS