432 Section on Health Policy Statistics A.M. Roundtable Discussion (fee event)

Section on Health Policy Statistics

Wednesday, August 3, 7:00 a.m.–8:15 a.m.

Preparing To Work In A Developing Country

✦Justin Fisher, U.S. Government Accountability Office, 441 G St., NW, Washington, DC, DC 20548, *fisherjs@gmail.com*

Key Words: developing nations, diversity, cultural sensitivity

Statisticians play an important role in international development. Working in developing nations presents both opportunities and challenges that should be considered before getting on the plane. During this session we will explore both the role of statistics and the soft skills (such as handling cultural diversity) involved in collaborative research projects with colleagues in developing nations. This session will be led by Justin Fisher, a statistician who has worked for several United Nations agencies (including UNCTAD, UNESCO and UNESCAP, the regional office for Asia and the Pacific) and is currently with the U.S. Government Accountability Office. As a member of Statistics Without Borders, he traveled to Haiti two months after the 2010 earthquake to assist with a survey designed to estimate the economic impact of the earthquake.

433 Section on Physical and Engineering Sciences A.M. Roundtable Discussion (fee event)

Section on Physical and Engineering Sciences

Wednesday, August 3, 7:00 a.m.–8:15 a.m.

Statistical Aspects Of Complex Computer Models

✦Dorin Drignei, Oakland University, , *drignei@oakland.edu*

Key Words: metamodels, sensitivity, calibration, computer experiment, surrogate, emulator

Computer models, or codes, are ubiquitous in science and engineering. In industry, computer models are often used to supplement expensive physical testing. There are inherent uncertainties associated with computer models, and statistics has already made significant contributions in this area. Still, much more needs to be done to address more complex computer models. During this roundtable we shall explore some of these complexities and how far can statistics go to address them. We hope to create a bridge between the academia and industry audiences, that will enhance further communication in this area.

434 Section on Statistical Consulting A.M. Roundtable Discussion (fee event)

Section on Statistical Consulting

Wednesday, August 3, 7:00 a.m.–8:15 a.m.

Deciding To Be A Statistical Consultant

✦Jesse Albert Canchola, Siemens Healthcare Diagnostics, 721 Potter Street, Berkeley, CA 94710, *jesse.canchola@siemens.com*; ✦Monica Johnston, MostlyMath, 2500 Camino Diablo, Suite 106, Walnut Creek, CA 94597, *MostlyMath@aol.com*

Key Words: consulting, business

Anytime one embarks on owning a Statistical Consulting, business there are many questions to consider. Does each question need to be fully examined or is there a set of key questions that one can address and then get down to business? What are the key administrative, financial, legal, operational, and personal factors one needs to consider? Does one go it alone or go into business with a partner? Would one first build a team of contractors or employees? Work out of a garage or pay for an office? Self-fund or try for a business loan? What basic legal and financial issues does one need to be aware of? What are the tax advantages? Statistical packages cost WHAT?!! Can one use open source software? Am I difficult to work with? Am I difficult to work for? Will I be happier professionally? Come with questions, come with answers.

435 Section on Statistical Education A.M. Roundtable (fee event)

Section on Statistical Education

Wednesday, August 3, 7:00 a.m.–8:15 a.m.

Developing Best Practices in Preparing Statistics PhDs to Teach

✦John Gabrosek, Grand Valley State University, 1 Campus Drive, Allendale, MI 49401-9403, *gabrosej@gvsu.edu*

Key Words: statistics education, academia, mentoring

Graduate programs in statistics do a commendable job of preparing PhD students to practice the discipline of statistics in industry, government, and academia. Unfortunately, most PhD programs do not adequately prepare a student to become a teacher of statistics at the university level. Students graduate without a working knowledge of best practices in statistics teaching and resources that are available for teaching statistics (particularly at the undergraduate level). This roundtable will be a brainstorming session for developing a set of best practices in mentoring graduate students who plan to teach and in mentoring new teachers of statistics at the university-level. The intended audience includes experienced teachers who have an interest in the preparation of graduate students who will become the next generation of statistics teachers.

Bringing Statistics Education To The Core

✦Nicholas Jon Horton, Smith College, Clark Science Center, 44 College Lane, Northampton, MA 01063-0001 USA, *nhorton@smith.edu*

Key Words: statistical education, statistical computing, modeling, Project MOSAIC, quantitative literacy

Do you believe that statistical thinking is at the core of a modern quantitative education? Project MOSAIC (www.mosaic-web.org) is a community of educators†working to create a broader approach to quantitative studies by integrating several important elements of quantitative training: Modeling, Statistics, Computation and Calculus. In this roundtable,†participants will learn about approaches to teach modeling and†computation in statistics, and discuss ways to integrate these activities†and approaches into their courses and curricula.

Developing The Next Generation Of Collaborators: How Can The Teaching Of Statistics Foster Positive Interactions With Researchers?

✦Deborah V Dawson, University of Iowa College of Dentistry, N439E Dental Science Building, U of Iowa College of Dentistry / Dows Institute, Iowa City, IA 52242-1010, *deborah-dawson@uiowa.edu*

Key Words: education, collaboration, research training

Bring any number of statisticians together and you are bound to hear a few "war stories" - tales of collaborations that were frustrating, suboptimal, or just plain nightmarish. How can we use the opportunities presented during our teaching to help young investigators learn to interact with statisticians in ways that facilitate the best research - and leave everyone satisfied - and even smiling? Bring your ideas and strategies to a discussion of how we can assist the next generation of researchers to be effective, informed collaborators.

Using Tinkerplots To Teach The Bootstrap Test In An Introductory Statistics Course

✦Rebekah Isaak, University of Minnesota, 56 East River Road, Minneapolis, MN 55455, *isaak009@umn.edu*

Key Words: modeling, simulation, bootstrap test, introductory statistics

The CATALST course is an introductory statistics curriculum that is built on ideas of modeling and simulation, with "the core logic of inference" as the foundation (Cobb, 2007, p. 13). When applied to randomized experiments and random samples, Cobb refers to this logic as the "three Rs": randomize, repeat, and reject. One activity presented in the second unit of this course uses the bootstrap test to model the variation due to random sampling under the assumption of no group differences. The focus is for the students to consider the study design when modeling the variation in the simulation. Students usemodeling and simulation software called TinkerPlotsto explore statistical concepts throughout the CATALST course, including the bootstrap test. Participants in this roundtable session will be givenexamples of activities, which includescreen shots from TinkerPlots.

436 Section on Teaching of Statistics in the Health Sciences A.M. Roundtable Discussion (fee event)

Section on Teaching of Statistics in the Health Sciences

Wednesday, August 3, 7:00 a.m.–8:15 a.m.

Information Criteria, Bayesian Methods And Their Implications For Statistical Literacy

✦Rochelle Elaine Tractenberg, georgetown university medical center, Neurology, Suite 202 Building D, 4000 reservoir Rd., NW, washington, DC 20007, *rochelle.tractenberg@gmail.com*

Key Words: statistical literacy, information criteria, bayesian methods, multi-model inference, null hypothesis significance testing

The use of information criteria (IC) or Bayesian methods (BM) for modeling, analysis, or decisionmaking, are treated as "advanced topics" and are at odds with null hypothesis significance testing (NHST). Can statistical literacy be achieved if NHST is the sole object of these literacy efforts? Options for integrating IC, BM, and NHST in statistical literacy curricula can reinforce current efforts, while not condemning the newly statistical literate to ignorance of these important advances in applied statistics, include: 1. pare down the NHST to accommodate incorporation of the other two topics. 2. choose IC or BM and integrate just that into the NHST training. 3. maintain the current emphasis on NHST in the development of statistical literacy, mentioning (leaving) IC and BM for advanced students. Considerations for revising/developing new statistical literacy programs include the preparation of NHST-oriented instructors to teach IC and BM, and the ramifications of limiting exposure to NHST and focusing on IC and/or BM. We will discuss these options, and others that participants may be contemplating for their own efforts to improve/achieve statistical literacy at their institutions.

437 Introductory Overview Lecture: An Overview of Current Problems in Statistical Genetics

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

An Overview of Current Problems in Statistical Genetics

✦Cavan Reilly, University of Minnesota, , *cavanr@biostat.umn.edu*

This presentation will provide a gentle introduction to statistical genetics and then present an overview of contemporary approaches and open problems. Topics will include, phenotype definition, population stratification, control of false positive rates, epistasis, gene-environment interactions, haplotype analysis and joint linkage and association analysis. Finally some novel approaches to detecting genetic variation related to disease, such as high throughput sequencing and exome scans, will be discussed.

438 Late Breaking Session: Supreme Court Finds Statistical Significance is not Necessary for Causation

ASA, ENAR, IMS, SSC, WNAR, International Chinese Statistical Association, International Indian Statistical Association

Wednesday, August 3, 8:30 a.m.–10:30 a.m.

Supreme Court Finds Statistical Significance is not Necessary for Causation

✦Steve Ziliak, Roosevelt University, , *sziliak@roosevelt.edu*; ✦Joseph ("Jay") Kadane, Carnegie Mellon University, Dept. of Statistics, 5000 Forbes Avenue, Pittsburgh, PA 15217, *kadane@stat.cmu.edu*; ✦Donald Rubin, Harvard University, Department of Statistics, 1 Oxford Street, Cambridge, MA 02138, *rubin@stat.harvard.edu*; ✦Daniel T Kaplan, Macalester College, 1600 Grand Ave., Saint Paul, MN 55105 USA, *kaplan@macalester.edu*

On March 22, the US Supreme Court found that Matrixx [as a public company] could not use the lack of statistical significance to withhold adverse information from investors."Matrixx's premise that statistical significance is the only reliable indication of causation is flawed." "Something more than the mere existence of adverse event reports is needed to satisfy that standard [material to a reasonable investor], but that something more is not limited to statistical significance." "Matrixx proposed bright-line rule requiring an allegation of statistical significance to establish a strong inference of scienter is once again flawed."This session examines the implications of the courts' finding for statistics in general and for statistical education in particular.

439 Lighting up the scoreboard: A look at statistical applications in basketball research ■

Section on Statistics in Sports

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Foul on the Committee? Investigating and Predicting NCAA Tournament Selection Committee Decisions

✦Jay Coleman, University of North Florida, Coggin College of Business, 1 UNF Drive, Jacksonville, FL 32224, *jcoleman@unf.edu*; Mike DeMond, Charles River Associates; Allen Lynch, Mercer University

Key Words: logit, probit, NCAA, basketball, group decisions, bias

The NCAA Men's Basketball Committee selects the 34 at-large teams for each season's tournament, and seeds the 65 participating teams. The great debate that follows the Committee's decisions each year often includes charges of bias in favor of major conferences and toward teams with some sort of Committee representation. Another common complaint is that the Committee seems to set limits on the number of teams that can be invited from any given league. In an extension of research on predicting at-large bids first published in 2001, we investigated the question of bias in the at-large selection of teams and the seeding decisions from 1999-2008, using logit and probit modeling. After controlling for team performance using 41 factors, we found evidence of bias in favor of nearly all major and even mid-major conferences in selection and/or seeding. We also found strong evidence of favoritism toward teams represented on the Committee. A variation of the at-large selection model predicted 32 of 34 teams selected by the Committee for the 2009 tournament, and 33 of 34 in 2010. Had the model's 2010 last-team-in and first-team-out predictions been reversed, it would have been perfect.

Life on the Bubble: Who's in and Who's out?

✦Scotland Charles Leman, Virginia Tech, Virginia Tech, 410 Hutcheson Bldg., Blacksburg, VA 24061 United States, *leman@vt.edu*

Key Words: March Madness, Bubble Teams, RPI, Strength of Schedule, Tournament, Basketball

Every year, the top National Collegiate Athletic Association (NCAA) basketball teams square off and compete for the national colligate basketball championship, in what is commonly referred to as "March Madness". However, before teams are eligible to compete, they must obtain entry into the tournament: either by winning their conference tournament, or by selection through a committee decision process. Various factors influence the committee's decision. Of primary importance is the perceived strength of a team; however, the possible measures used to assess a team's strength can vary widely. For teams, on the borderline of tournament inclusion, these metrics are of particular interest. Using historical data, we assess which metrics play the most substantial roles in predicting tournament inclusion. Based on these metrics, we determine guidelines for constructing a team's schedule, in oder to enhance the chance of being selected for the tournament. We conclude by discussing alternative metrics, and procedures which may be used to obtain closer agreement between currently contrasting metrics.

Shooting for the Top Spot: Exploring the Use of NBA Combine Data to Predict Draft Position

✦Clayton Barker, SAS Institute, 100 SAS Campus Drive, Cary, NC 27513, *clay.barker@sas.com*; Michael T Crotty, SAS Institute

Key Words: basketball, NBA draft, ranking, variable selection, data mining, logistic regression

Each summer, the National Basketball Association (NBA) drafts sixty college and international players into the league. The teams draft sequentially, so it is crucial to choose the "best" player among the remaining players. In order to help teams make these decisions, the NBA combine is an event held before the draft that measures various physical and athletic ability attributes of each player. Based on these measurements and college or international performance, NBA teams choose the player that they expect to make the greatest contribution to their team. The data from the combine are often cited during pre-draft analysis and predictions, but what effect, if any, do the combine measurements have on player selection? Using a combination of variable selection and data mining techniques, we will determine which factors impact draft position and we will attempt to predict the 2011 NBA draft. We also compare the predictability of the draft using the combine data to using the college and college conference affiliation of the player.

Ranking Teams in the NCAA Division I Men's Basketball Tournament

✦Mark Brown, City College, CUNY, Mathematics Department, City College, CUNY, NY, NY 10031, *cybergarf@aol.com*; Joel Sokol, Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology

Key Words: Empirical Bayes, Markov chains, Prediction, Quantitative sports analysis, Applied probability, Applied statistics

The LMRC method for predicting NCAA tornament results from regular season game outcomes is a two part process. First using a logistic regression model we focus on those pairs of teams who have played one another during the regular season. For these pairs we estimate the probability of one team beating another on a neutral court. These probabilities are utilized in a probability transition matrix of a Markov chain. All teams are then ranked by the stationary distribution of this chain. This provides methodology to rank teams who have not played one another during the regular season, and in some cases have no common opponents either. In recent work, (Brown, M. and Sokol, J. (2010), Journal of Quantitative Analysis in Sports, Volume6, Issue 3, Article 4), we replaced the logistic regression analysis by an empirical Bayes approach. This resulted in significant improvement both on past tournament data and in the 2010 tournament in which the method correctly predicted the winning team in more tournament games than any of the few dozen competing methods found online. In this talk the methodology will be reviewed and we will report on its performance in the 2011 tournament.

Building a Championship Contender: A Statistical Model for NBA Lineup Combinations

✦Ryan J. Parker, North Carolina State University, 1342 Rio Valley Dr, Apt 206, Raleigh, NC 27614, *rjparker@ncsu.edu*; Brian Reich, North Carolina State University; Jarad Niemi, University of California, Santa Barbara

Key Words: basketball, lineup combinations, team performance

Several methods have been proposed to evaluate the contribution of individual basketball players to their team's success. In this talk we study the performance of combinations of NBA basketball players. In basketball, like many other popular American sports, team performance is not be well-predicted by the sum of the performance of the team members. For example, it may be that a three-point shooter's value is increased when paired with a play-making point guard or a post player that draws double teams. A challenging part of putting together an effective roster of players is accounting for these synergetic relationships while extrapolating to unobserved combinations of players. Combining individual player statistics and results for all 10-player combinations for the 2009-2010 NBA season, we build a statistical model for the performance of a 10-player combination that accounts for potentially-complex interactions between types of players, and we show that this model outperforms other proposed predictive models that simply use the average ratings of the players. Further, we demonstrate how to use our approach for free agent signing and setting lineups for particular opponent matchups.

440 Better Medical Decision-Making through Dynamic Visualization of Benefit-Risk ■●

Biopharmaceutical Section, Section on Statistical Graphics, Section on Risk Analysis, Section for Statistical Programmers and Analysts

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Seeing Is Believing: Graphical Depictions of Low-Probability Outcomes for Eliciting Treatment-Related, Benefit-Risk Tradeoff Preferences

✦Reed Johnson, Research Triangle Institute, PO Box 12194, Research Triangle Park, NC 27709 USA, *frjohnson@rti.org*; Lauren Donnalley, Research Triangle Institute; Juan Marcos Gonzalez, Research Triangle Institute

Key Words: Numeracy, Benefit-Risk Analysis, Health Risk Communication, Visual Aids

Individuals with low numeracy skills have been shown to have different perceptions of presented risk then those with high numeracy skills. Effective risk communication is essential and challenging when asking subjects to evaluate treatment benefits and risks. Researchers have proposed visual aids for communicating quantitative risks; however, empirical evidence is sparse about which format is most effective for specific judgment tasks. This study evaluated graphical depictions of medical-outcome probabilities from 19 otherwise similar discrete-choice experiments conducted in 9 different countries (n=8,111) to evaluate how graphic format affected understanding of numerical probabilistic risks. Three risk-grid variations were tested with either 100 or 1000 units. Random-effects panel regression analysis of correct interpretation of the graphic parameters yields negative and significant parameters for age and not being a US subject, and positive and significant parameters for education and humanoid icons compared to square or dot icons. Better comprehension of quantitative risks also is shown to improve the statistical precision of preference-parameter estimates.

Dynamic Visualization Tools to Simplify the Analysis of Complex Clinical Drug Safety Data

✦Ana Szarfman, FDA, , *ana.szarfman@fda.hhs.gov*

Key Words: interactive graphic display, drug safety, automated analyses, patient profiles, drill down to source data, data auditing

Interactive, automated data visualization tools provide disciplined approaches to investigate and document potential drug safety problems within different data resources. These tools display the sequence of events for multiple patients and within single patients allowing the user to quickly switch from one modality to the other. They reveal the data peculiarities and missing data. To understand the global picture, rows of multi-patient data are aligned by start of treatment or other event and sorted by duration of treatment or time to an outcome. To understand individual patients, multiple data types are put on a common timeline, aligned by start time of a drug or an event and linked to narrative text. This display also shows the data points and the start and end dates of medications, events, indications, and procedures. The ability to "drill down" from a global display or from any result to the individual patient profile and source data increases the certainty about the results. With automation, multiple reviewers can use the same datasets and outputs, so that the analytical processes can be easily understood and reused, fostering communication between all parties.

Supporting Medical Decisionmaking Through Dynamic 3D Visualization of Risk, Benefit, and Uncertainty

◆Richard A Forshee, U.S. Food and Drug Administration, CBER, 1401 Rockville Pike, HFM 210, Rockville, 20852, *Richard.Forshee@fda.hhs.gov*; Arianna Simonetti, U.S. Food and Drug Administration, CBER; Mark O Walderhaug, U.S. Food and Drug Administration; Anne Fernando, Norfolk State University

Key Words: Risk-Benefit assessment, visualization, medical decision-making, biologics, vaccine

One of the challenges for Risk:Benefit analysis is to develop tools to help decision-makers analyze the risks and benefits of products while properly characterizing the uncertainty of the estimates of risks and benefits. We have developed a three-dimensional visualization tool where the height of the surface represents the probability that the treatment has a particular risk-benefit coordinate. The model is applied to historical data from clinical trials and post-market surveillance for a vaccine that was determined to have an elevated risk for an adverse event. The tool shows key properties of distributions that may affect decision-making; integrates data on multiple properties of a medical product; demonstrates the evolution of scientific knowledge, and encourages stakeholders to interact with the visualization to explore what-if scenarios.

A Graphical Method for Patient-Level Benefit-Risk of a Medical Product Over Time, with Examples

◆Jonathan D Norton, Center for Drug Evaluation and Research, US Food and Drug Admin., 10903 New Hampshire Avenue, WO22-3643, Silver Spring, MD 20993, *Jonathan.Norton@fda.hhs.gov*

Key Words: benefit-risk, risk-benefit, graphics, safety, longitudinal, missing data

Both FDA and external stakeholders acknowledge a need for improved benefit-risk assessment. This problem is challenging for a number of reasons: benefits and harms tend to be measured on different scales, they may be statistically associated, their relative balance may change over time, and there may be substantial missing data. The Individual Response Profile, a recently-developed method for simultaneously visualizing benefit and risk over time, is presented. This method addresses all of the aforementioned issues, particularly in regard to treatments that provide only symptomatic relief. The underlying model represents a subject's benefit-risk at a given time as one of five discrete clinical states, one being premature study withdrawal. The graphic uses colors to represent each subject's changing state over the course of the clinical trial. The user can quickly grasp how a treatment affects subjects in aggregate, and then further examine how individuals are affected. It is possible to tell whether the beneficial and harmful outcomes are correlated. A chronic pain trial and other worked examples are presented, and inferential approaches are discussed.

441 Design and Analysis Strategies for Efficient Interaction Modeling in Epidemiologic Studies ■✪

Section on Statistics in Epidemiology, ENAR, International Indian Statistical Association

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

A Bayesian Model Averaging Approach to Gene-Environment Interaction in a GWAS

Dalin Li, University of Southern California; ◆David V. Conti, University of Southern California, , *dconti@usc.edu*

Key Words: Gene-Environment Interaction, Genetic Association, Bayesian Model Averaging

Scans for gene(G)-environment(E) interaction have been mostly neglected in current GWAS. The conventional approach, the case-control analysis, suffers from low power. The alternative case-only analysis can be more powerful but violation of the assumption of independence can greatly bias the results. We propose a model averaging approach for GxE interaction leveraging the fact that most SNPs in a GWAS are not associated with either the outcome nor the environmental factor. When there is no G-E association or main effect of G, removing the corresponding terms in the log-linear model reduces the variance of the interaction estimate. Since there is uncertainty, the method averages over the sub-models with and without the G-E or main effect G terms. A final estimate of effect is the average over the models weighted by the posterior probability of each model. In a single-marker analysis this approach can be more powerful than the case-only approach when the GxE independence assumption holds. When the assumption is violated, the Type I error rate is more robust than the case-only analysis. In genome-wide simulations, this approach has better overall performance.

Assessing GxE Using Nuclear Families: Can We Make It Robust?

◆Clarice Weinberg, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709 USA, *weinber2@niehs.nih.gov*; Min Shi, National Institute of Environmental Health Sciences; David Umbach, National Institute of Environmental Health Sciences

Key Words: case-only, family-based analysis, gene-by-environment interaction, environment

Complex diseases arise through joint action of environmental factors and genetic susceptibility variants. Investigators have assumed that the robustness protecting family-based genotype analyses from bias due to population stratification will protect assessments of interaction. Not true. When testing against a multiplicative joint effects model, the case-only design offers good power, but is invalid if G and E are correlated in the source population. Four mechanisms can produce G/E dependence: 1. Subpopulations vary in both ancestry and cultural traditions; 2. Family history influences behavior through avoidance of exposures perceived as risky; 3. G influences E; or 4. Selective attrition produces correlation in older age strata. We propose a sibling-augmented case-only approach, which is robust against the first two mechanisms, and is therefore valid when a young-onset disease is studied and G does

not influence E. With this design an unaffected sibling is randomly selected and exposure ascertained for that sibling. The case is genotyped if the case-control pair is exposure-discordant. A logistic model permits testing of the exposure main effect and GxE interaction.

Interaction Assessment in Moderately Sized Randomized Studies

✦Michael LeBlanc, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109 USA, *mleblanc@fhcrc.org*

Key Words: subgroups, interactions, regularization

We develop efficient modelling strategies for genetically defined subject subgroups. These methods build on a special case of general interaction models which are sometimes referred to as varying coefficient or effect models. The models are of the form M(T)[(G)] representing a treatment model that potentially varies with respect to genomic profiles or "gating" of the treatment effects or toxicity associations. We study the impact on inferences when T is defined by randomized treatment assignment and explore the connection to other observational genetic association studies where the genetic association can be thought to be "gated" by certain environmental or risk behaviour variables. To facilitate use of the results for future prevention and therapeutic trial design, the statistical method we develop will allow investigator control over class of subset rules and the fraction of subjects identified by the rules. We extend the class of indexed rules originally developed for simple prognostic models in the Extreme Regression method. In addition, we introduce additional variance control and variable selection for subgroups via Elastic-Net type regularization.

Efficient Two-Step Testing of Gene-Environment and Gene-Gene Interactions in Genome-Wide Association Studies

✦Cassandra Elizabeth Murcray, University of Southern California, 1540 Alcazar Street, CHP 222Q, Los Angeles, CA 90089, *murcray@usc.edu*; Duncan Thomas, University of Southern California; Juan Pablo Lewinger, University of Southern California; W. James Gauderman, University of Southern California

Key Words: GWAS, gene-gene interaction, gene-environment interaction, case-control, case-parent trios

Recently developed methods to detect interactions in GWA studies have shown increased power relative to traditional approaches. Two-step analyses have been proposed to prioritize the large number of SNPs tested to highlight those likely to be involved in GxE/GxG interactions. Kooperberg and LeBlanc (2008) suggested screening on genetic marginal effects in a search for GxG interactions. Alternatively, Murcray et al (2009) suggested screening SNPs by testing the G-E association induced by an interaction in the combined case-control sample. Gauderman et al (2010) proposed a screening step based on the association between parental genotypes and case exposure in GWA studies in case-parent trios. In these methods, SNPs that pass the respective screening step at a liberal significance threshold are followed up with a formal test of interaction in the second step. For case-control data, we show that the Murcray et al approach is often the most efficient method, but that a hybrid method that combines the Murcray et al and Kooperberg et al methods by allocating a proportion of the experiment-wise significance level to each test is a powerful and robust method for nearly any underlying model.

442 Defense and National Security Applications of Text Mining ■●

Section on Statistics in Defense and National Security, Section on Government Statistics, Section on Statistical Computing, Section on Statistical Graphics, Section on Physical and Engineering Sciences, Section on Risk Analysis, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Joint Embedding of Disparate Information with Applications in Text Analysis

✦David Marchette, Naval Surface Warfare Center, , *dmarchette@gmail.com*; Carey Priebe, Johns Hopkins University

Key Words: multidimensional scaling, disparate data fusion, dimensionality reduction, text data mining, implicit translation

The analysis of scientific documents often results in the extraction of features of very different type. For example, in addition to standard text analysis features such as used in bag-of-words models, features extracted from natural language processing methods, or similar text-based features, one may extract graphs such as co-authorship or citation networks, features extracted from images or figures within the documents. One way to perform inference on such different and complex data is to embed the data into a single space so that the different types of information can be combined in a useful way. We discuss several approaches to this problem, derived from a dissimilarity framework and utilizing multidimensional scaling ideas to define the embeddings. We will illustrate these techniques on several text document datasets.

Estimation and Fusion of Multidimensional Cluster Signatures

✦Jeffrey Solka, NSWCDD, , *jeffrey.solka@navy.mil*

Key Words: cluster analysis, visualization, fusion, horizon scanning, text analysis, information exploitation

This talk will discuss our recent efforts in the identification of interesting clusters within a large collection of clusters. Given a large corpus of documents our research has focused on the identification of interesting clusters. The concept of interest is often vague being user and application dependent. Interest might be related to content in a particular area or rate of growth of the number of publications within a particular cluster. This talk will discuss our strategy for clustering large document collections, some of the signatures that one might wish to calculate on these clusters, and various schemes to fuse these signatures together to produce a sorted list of clusters for human inspection. Anecdotal results will be presented on a large set of PubMed abstracts.

Computing Within the Foreign Vector Space Framework

✦Nick Tucey, Naval Surface Warfare Center, , *nicholas.tucey@navy.mil*; Avory Bryant, Naval Surface Warfare Center; Kristen Ash, Naval Surface Warface Center

Key Words: term document matrix, information retrieval, foreign language text processing, multilingual documents

This talk will discuss how to expand a text data mining application to allow for the processing of documents in foreign languages, such as Arabic, Chinese, Russian, etc. Many text analysis algorithms, such as hierarchical clustering, latent semantic indexing (LSI), and dimensionality reduction, utilize a vector space encoding of the document collection, and thus are language independent. These techniques, however, are all dependent on the tokenization of multilingual documents. This talk will explore topics such as language identification, word segmentation, indexing, and document display. In order to develop this methodology necessary and sufficient conditions for each method are studied. After that, maximum likelihood estimators for the parameters are found. This allows us to develop inference procedures that are shown to work in simulation studies. The models are applied to get conclusions in real data set examples. Finally, its connection with principal components is given.

443 Statistics in Business ■●

IMS

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Optimal Parameter Estimation and Portfolio Choice: A Survey and New Results

✦Guofu Zhou, Washington University, Olin School of Business, Washington University, St. Louis, MO 63130 United States, *zhou@wustl.edu*

Key Words: optimal estimation, shrinakge, portfolio choice, Bayesian

In this talk, we provide a survey of the literature on mean-variance portfolio choice that reply on estimated mean and covariance matrix of asset returns, with both recent and new developments. Our survey covers both classical and Bayesian methods. The classical approach uses shrinkage, combination and approximate models to yield better estimates, while the Bayesian strategy incorporates information about events, macro conditions, asset pricing theories, and security-driving forces to form useful priors in selecting optimal portfolios. We also review a range of applications, and outline some challenges for future research.

Modeling Topic Selection of Web Browsing Using Clickstream Data

✦Alan Montgomery, Carnegie Mellon University, 5000 Forbes Ave., Tepper School of Business, Pittsburgh, PA 15213 USA, *alanmontgomery@cmu.edu*

Key Words: Bayesian Analysis, Topic Models, Clickstream Data, Consumer Behavior, Marketing

Users appear to view websites in sequences that are related to a latent topic. For example, a user may have a session that is focused on gathering information about a product, which may have a large number of viewings at promotional, corporate, and portal sites, while news gathering may have a very different profile. The idea is that consumers group their activities together into related topics. The goal of this study is to detect the underlying topics that are driving user browsing behavior using a correlated topic model. Specifically we use a multivariate normal to model the log-odds ratio of a given topic being chosen. The correlations permit relationships amongst the topics. Conditional upon the topic each website is chosen using a choice model across all websites. Our model is related to a latent Dirichlet allocation and correlated topic models employed in text analysis. We consider generalizations with dynamic trends to understand how topic selections may depend upon time.

Extreme Value Theory in the Study of Ruin with Risky Investments

✦Qihe Tang, The University of Iowa, Department of Statistics and Actuarial Science, 241 Schaeffer Hall, Iowa City, IA 52242-1409, *qihe-tang@uiowa.edu*; Raluca Vernic, Ovidius University of Constanta; Zhongyi Yuan, University of Iowa

Key Words: Asymptotics, Finite-time ruin probability, Max-domain of attraction, Multivariate regular variation, Optimal investment strategy, Subexponential distribution

Consider a discrete-time insurance risk model in which periodic claim amounts and premium incomes form a sequence of independent and identically distributed random pairs, each with dependent components. The insurer is allowed to invest a constant fraction of his/her wealth in a risky stock and keep the remaining wealth in a risk-free bond. For subexponential periodic net losses and arbitrarily dependent stock return rates, we derive a general exact asymptotic formula for the finite-time ruin probability. If the loss distribution also belongs to the max-domain of attraction of the Fr\'{e}chet or Gumbel distribution and the return rates jointly follow Sarmanov's distribution, the obtained general formula is further refined to be completely transparent. As an application, we approximate the value of the fraction invested to the risky stock that maximizes the expected terminal wealth of the insurer subject to a constraint on the ruin probability.

Expectile Regression with Varying Coefficient

✦Yong Zhou, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, No.55, Zhongguanchun East Road, Haidian District, Beijing, 100190 P.R. China, *yzhou@amss.ac.cn*

Key Words: Expectile regression, Asymmetric least square, Varying-coefficient model, Local linear regression

In this paper, we consider a regression expectile with varying-coefficient function used to characterize the relationship between a response variable and explanatory variable when the behavior of "non-average" (or extreme behavior) individuals is of interest. Regression expectile is an alternative location measure of conditional distribution defined by an asymmetric least squares criterion function to regression quantile. The expectile estimators have properties which are analogous to regression quantile estimators, but are much simpler to compute by iteratively reweighted asymmetric least squares. In view of the appealing algorithm of the expectile and the link with VaR (value at risk) and ES (expected shortfall), which are popular measures of financial risk, expectile can also to be used to estimate VaR and ES in an indirect way.

444 Recent development in Causal inference ■●

ENAR, International Indian Statistical Association, Section on Health Policy Statistics, Section on Statistics in Epidemiology, Section for Statistical Programmers and Analysts

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Using the Instrumental Propensity Score in Observational Studies for Causal Effects

✦Jing Cheng, UCSF School of Dentistry, 3333 California St, Ste 495, San Francisco, CA 94143 USA, *jing.cheng@ucsf.edu*

Key Words: Causal effect, instrumental variable, instrumental propensity score, obersvational studies

In observational studies, investigators are often concerned about unmeasured confounding in the evaluation of the effect of a program/exposure/treatment. The instrumental variable approach is valuable to address unmeasured confounding and is desired in many applications. However, many existing IV methods require the IV to be independent of all measured and unmeasured confounders and investigators often have difficulty to find a good IV. In this study, we propose to use instrumental propensity score to relax the assumptions for an IV to be valid and then use subclassification or weighting to conduct inference on causal effects.

The Importance of Underlying Ordering Assumptions in Causal Inference

✦Jessica Gerald Young, Harvard School of Public Health, 677 Huntington Avenue, Kresge, 820, Boston, MA 02115, *jyoung@hsph.harvard.edu*; Miguel A. Hern·n, Harvard School of Public Health; James M. Robins, Harvard School of PUblic Health

Key Words: causal inference, assumptions, ordering, parametric g-formula, inverse probability weighting, marginal structural models

When estimating the causal effect of a time-varying treatment using longitudinal data, most applications assume that the ordering of confounders with respect to treatment is known. In interval cohorts, however, the ordering of covariates measured simultaneously is generally unknown. While much attention has been paid to the importance of assumptions in making causal inferences, the consequences of assumptions concerning ordering are often ignored. Effect estimates may be sensitive to the choice of ordering and investigators must be careful in defining their estimator, as well as selecting an appropriate estimation algorithm, depending on this choice. Here we consider various underlying ordering assumptions and how they impact effect estimators and associated estimation algorithms under interventions on single and multiple treatments. Focus is on estimators based on the parametric g-formula and inverse probability weighting of marginal structural models. Examples are given from the coronary heart disease and HIV literature.

Methods for Clustered Encouragement Design Studies with Missing Data

✦Leslie Taylor, VA Puget Sound Health Services R&D, 1100 Olive Way, Suite 1400, Seattle, WA 98101, *taylorl@u.washington.edu*; Xiao-Hua Andrew Zhou, Department of Biostatistics, University of Washington

Key Words: causal inference, clustered encouragement design, complier average causal effect, missing data, multiple imputation, principal stratification

Encouragement design studies are particularly useful for estimating the effect of an intervention that cannot itself be randomly administered to some and not to others. They require a randomly selected group receive extra encouragement to undertake the treatment of interest, where the encouragement typically takes the form of additional information or incentives. We consider a “clustered encouragement design” (CED), where the randomization is at the level of the clusters (e.g. physicians), but the compliance with assignment is at the level of the units (e.g. patients) within clusters. Noncompliance and missing data are particular problems in encouragement design studies, where encouragement to take the treatment, rather than the treatment itself, is randomized. The motivating study looks at whether computer-based care suggestions can improve patient outcomes in veterans with chronic heart failure. Since physician adherence has been inadequate, the original study focused on methods to improve physician adherence, although an equally important question is whether physician adherence improves patient outcomes. Here, we reanalyze the data to determine the effect of physician adherence.

445 Well-known graphics software that is little known by statisticians ■●

Section on Statistical Graphics, Section on Statistical Computing, Section for Statistical Programmers and Analysts, International Indian Statistical Association

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Tableau Software: Finding and Telling Stories with Data

✦Jock D. Mackinlay, Tableau Software, 837 North 34th Street, Suite 400, Seattle, WA 98103, *jmackinlay@tableausoftware.com*

Key Words: visualization, interaction, user interfaces, visual analysis, databases, storytelling

In recent years, interactive graphical visualization has begun to move from research labs into widespread use. Our vision is that everyone will be able to find stories in their data and tell their stories on the web with interactive visualizations as easily as they now share images and video. To that end, we have developed an algebra that describes visualizations of data including tables, charts, graphics, maps, time series and small-multiple views of data. The algebra is the foundation of an intuitive drag-and-drop interface for rapidly finding stories in data. This interface is in widespread use. The algebra also has the important property that it compiles into database queries, which connects the visual interface to most database technology including very large databases. Users can also use this interface to author interactive applications to the web to tell stories with their data. We are developing a platform for storytelling with data that scales to large data sets and the audience sizes of the web.

Spotfire: Interactive Visualizations to Complement Statistical Analysis

✦Stephen Kaluzny, TIBCO Software Inc, 1700 Westlake N, #500, Seattle, WA 98109, *skaluzny@tibco.com*

Key Words: visualization, R, Splus

We will provide an overview of Spotfire, focusing on how Spotfire's interactive visualizations complement and enhance the statistical analysis of data. Spotfire's interactive graphics provide features such as dynamic drill-down and interactive filtering to quickly focus on and analyze subsets of data. Its visualizations can be used side-by-side with statistical software for ad hoc analysis and exploration, or integrated directly with analyses in S+ and R to create relevant, powerful analytic applications for use by a wide audience without advanced statistical training.

Protovis: A Declarative Language for Interactive Visualization

✦Jeffrey Heer, Stanford University, CA 94305-9035, *jheer@cs.stanford.edu*

Key Words: visualization, design, toolkits, user interfaces, interaction

Despite myriad tools for visualizing data, there remains a gap between the notational efficiency of high-level visualization systems and the expressiveness and accessibility of low-level graphical systems. Powerful visualization systems may be inflexible or impose abstractions foreign to visual thinking, while graphical systems such as rendering APIs and vector-based drawing programs are tedious for complex work. We argue that an easy-to-use graphical system tailored for visualization is needed. In this talk, we will present Protovis, an extensible toolkit for constructing visualizations by composing simple graphical primitives. In Protovis, designers specify visualizations as a hierarchy of marks with visual properties defined as functions of data. This representation provides a level of expressiveness comparable to low-level graphics systems, while improving efficiency--the effort required to specify a visualization--and accessibility--the effort required to learn and modify the representation. We substantiate this claim through a diverse collection of examples and comparative analysis with popular visualization tools.

446 What We Know So Far About the 2010 Census - Operational Outcomes ■

Social Statistics Section, Section on Government Statistics, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

The 2010 Census Communications Campaign

✦Burton Reist, U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746, *burton.h.reist@census.gov*; Megan Kindelan, U.S. Census Bureau

Key Words: 2010 Census, Communications

As 2010 Census planning began, we faced challenges such as an economic crisis; increased diversity; a rapidly changing communications environment; and increased government distrust. Strategically, we knew overcoming these challenges would require sophisticated audience segmentation research and flexible, integrated communications approaches to surround audiences with census messages. Why was the integrated communications campaign necessary? It comes down to a business decision-spend money upfront to inform and motivate the public to mail back their form and, in turn, save millions on the overall cost of the 2010 Census. A number of strategies and tactics will be explored to determine how well our efforts worked and the lessons learned. For example, did our in-language materials/advertising allow us to effectively communicate to hard-to-count audiences? And did our participation rate monitoring tools give us timely feedback so we could adjust our communications strategies and mobilize the hardest to count in real time? The 2010 Census ultimately matched the 2000 participation rate. This paper will probe what impact the communications campaign had on this success.

The Effect of Programs to Increase Response in Census 2010

✦Robin Amie Pennington, U.S. Census Bureau, , *robin.a.pennington@census.gov*

Key Words: participation rates, bilingual questionnaires, replacement questionnaires

Following up on non-respondents is one of the most costly aspects of any survey. On the scale of the decennial census, the cost of obtaining household data from these non-respondents through field work becomes astronomical. Many methods for increasing the response rate were instituted into Census 2010. The most visible were publicity and advertising efforts, including Census Bureau website postings of daily participation rates at various geographic levels. There were also two major methodological changes to the census aimed at increasing response. In particular, English/Spanish questionnaires were delivered in areas where data indicated a concentration of housing units where at least one adult in the household spoke Spanish and did not speak English very well. In addition, replacement questionnaires were mailed in areas where low response rates were expected. In this paper we will discuss Census 2010 participation and response rates, as well as some results on the impact of our use of the English/Spanish questionnaire and the replacement questionnaire.

Field Enumeration Outcomes in the 2010 Census

✦Maryann McCormick Chapin , U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *maryann.m.chapin@census.gov*; Annetta Clark Smith, U.S. Census Bureau

Key Words: Nonresponse Follow-up, Coverage, Group Quarters, Enumeration

A goal of the 2010 Census was to count people once, only once, and in the right place. Although the Census Bureau relies heavily on the public returning responses by mail, that alone is not sufficient to ensure everyone has an opportunity to be counted. The Census Bureau uses multiple enumeration methods in an attempt to reach its goal. These enumeration methods include: Update/Enumerate, Update/Leave, Group Quarters Enumeration, Remote Alaska, and Nonresponse Follow-up. Additional opportunities, such as Telephone Questionnaire Assistance and the Be Counted Program, are provided if an individual or household thinks they have been missed. Coverage operations such as Field Verification and the Coverage Follow-up provide a final opportunity to ensure accurate coverage in the Census. This paper presents highlights from these 2010 Census enumeration operations. It reviews the general methodology, implementation challenges, results of the enumeration, and key lessons learned from these operations.

447 Medallion Lecture: Connecting Experimental Data on Genes with Functional Information ■

IMS, International Chinese Statistical Association, International Indian Statistical Association, SSC

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Medallion Lecture: Connecting Experimental Data on Genes with Functional Information

✦Michael A Newton, University of Wisconsin, Madison, , *newton@biostat.wisc.edu*

Key Words: genomics, data integration, pleiotropy, gene ontology, graphical model

A problem in statistical genomics is to examine the points of contact between genomic data generated experimentally and exogenous information about gene function. The purpose of such data integration may be to summarize extensive gene-level data into manageable units, or it may be to enhance the signal to noise ratio through set-level averaging. In either case there are unique statistical problems with such data integration. I will review several statistical approaches and examine them in examples from cancer virology and flu replication genomics. Included will be a discussion of a new ``role model', which aims to address pleiotropy and the spurious association of gene sets with changes in cellular state. I will assume minimal knowledge of biology, and will aim to focus my presentation on the interesting statistical elements.

448 Statistics - the secret weapon of successful web giants ●

Section on Statistics and Marketing, International Chinese Statistical Association, Section on Statistical Computing

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Conditional Regression Models

✦William D Heavlin, Google, Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043, *bheavlin@google.com*

Key Words: boosting, causal inference, likelihood, machine learning, noise factors, von Mises-Fisher distribution

Because of the large volumes of data available, many web traffic investigations progress successfully using only simple two-sample experiments. However, an important subset of studies -- those dealing with rare events, causal inference, and/or measurement validation -- benefit from fine-grained blocking. By forming likelihoods free of the block nuisance parameters, conditional models point to a mathematically compact, parsimonious approach. Further, the conditional Gaussian, conditional Poisson, and conditional logistic models all share a single functional form. This talk discusses why one might want apply conditional models, how to estimate their coefficients and compute them efficiently, and how to adapt and deploy them as predictive models. We illustrate with a variety of Google examples.

The Effectiveness of Display Ads

✦Tim Hesterberg, Google, Inc., 651 N. 34th Street, Seattle, WA 98103 United States, *rocket@google.com*; Diane Lambert, Google; David X. Chan, Google; Or Gershony, Google; Rong Ge, Google

Key Words: ad effectiveness, display ads, causal modeling, irrelevant outcomes, observational data, selection bias

Display ads proliferate on the web, but are they effective? Or are they irrelevant in light of all the other advertising that people see? We describe a way to answer these questions, quickly and accurately, without randomized experiments, surveys, focus groups or expert data analysts. Causal modeling protects against the selection bias that is inherent in observational data, and a nonparametric test that is based on decoy outcomes provides further defense. †Computations are fast enough that all processing, from data retrieval through estimation, testing, validation and report generation, proceeds in an automated pipeline, without anyone needing to see the raw data.

Measuring Ad Effectiveness Using Continuous Geo Experiments

✦Jon Vaver, Google, Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043, *jvaver@google.com*; Deepak Kumar, Google, Inc.; Jim Koehler, Google, Inc.

Key Words: geo experiments, continuous measurement, online advertising, AdWords, Google

Google uses geo experiments to help AdWords advertisers assess the effectiveness of their paid search advertising spend. These experiments measure the value of incremental ad spend, identify optimal bid levels, quantify the impact of online ad spend on offlines sales, etc. In these experiments, geos are randomly assigned to a control or treatment group and users in these geos are served, or not served, search ads accordingly. Typically, such one-off experiments require a custom setup, they are disruptive to the advertiser's marketing plan, and the results only reflect the competitive landscape at a snapshot in time. In this paper, we describe the adaptation of geo experiments to a framework of continuous measurement. This approach mitigates the issues stated above while providing more definitive results than equivalent one-off experiments. For example, results indicate that a continuous experiment that changes spend in 1/8th of the geos can achieve the same confidence interval as an analogous one-off (non-continuous) experiment that changes ad spend in 1/2 of the geos.

When Less Is More and More Is More: Proximity and Repetition in Online Display Retail Advertising in a Large-Scale Field Experiment on Yahoo!

✦Randall A. Lewis, Yahoo! Research, 4401 Great America Parkway, 2GA-2616, Santa Clara, CA 95054, *ralewis@yahoo-inc.com*; David H. Reiley, Yahoo! Research; Garrett A. Johnson, Northwestern University Economics

Key Words: field experiment, online display advertising, frequency, distance, offline sales, advertising effectiveness

Yahoo! Research partnered with a nationwide retailer to study the effectiveness of display advertising on online and in-store sales on more than three million shared customers. We measure the impact of higher ad impression frequency using a simple design that varies the ads that users see on the Yahoo! network within identically targeted campaigns: users in the treatment group see the retailer's ads; users in the control group see unrelated, `control' ads; and users in the half treatment group see an equal probability mixture of the retailer's and control ads. We find a statistically significant increase in sales, relative to the control group, as a result of the ads. Doubling the number of impressions per person--from 17 to 34 in a two-week period--approximately doubled the treatment effect. We also find striking evidence that the ads most strongly affected customers who live closest to the retailer's brick-and-mortar locations: those who live within two miles of a store experience six times the incremental sales lift due to ads as those living farther away.

449 Celebrating Wayne Fuller's 80th birthday ●

General Methodology, International Chinese Statistical Association, International Indian Statistical Association, SSC

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Impact of Wayne Fuller's Contributions to Sample Survey Theory and Practice

✦J. N.K. Rao, Carleton University, School of Mathematics and Statistics, Carleton University, Ottawa, ON 2G 4H8 Canada, *jrao34@rogers.com*

Key Words: sample surveys, regression analysis, small area estimation, imputation, two-phase sampling, measurement errors

Wayne Fuller has made fundamental contributions to three major areas in statistics, sample surveys, time series analysis and measurement errors, often interfacing between them. In this talk I will focus on his major contributions to sample survey theory and practice, covering a wide range of topics including regression estimation and analysis, measurement errors, small area estimation, two-phase sampling, imputation for missing data and software for survey data analysis. I will highlight the impact of his work on survey practice and development of sampling theory.

Model-Based and Semiparametric Estimation of Time Series Components and Mean Square Error of Estimators

✦Danny Pfeffermann, University of Southampton and Hebrew University, Southampton Statistical Sciences Research Institut, University of Southampton, Southampton, SO17 1BJ United Kingdom, *msdanny@huji.ac.il*; Richard Tiller, Bureau of Labor Statistics; Michail Sverchkov, Bureau of Labor Statistics

Key Words: Bias corrections, Mean-Square error, Seasonal Adjustment, Trend

Wayne Fuller is known for his outstanding contributions to three main areas in statistics: Sample survey theory, Time series analysis and Measurements errors. This presentation will focus on time-series analysis and more specifically, on estimation of seasonally adjusted and trend components and the MSEs of the estimators. We shall compare the component estimators as obtained by application of the Tramo-Seat procedure, X-12 ARIMA and by fitting State-Space models that account more directly for correlated sampling errors. The classical model-based component estimators and variance estimators will be compared with a new procedure, which uses a different definition of the components that conforms to an original suggestion by Wayne Fuller. By this definition the unknown components are defined to be the hypothetical X-11 estimates if sufficient data was available for application of the symmetric filters imbedded in this procedure. The new MSE estimators are with respect to this definition. The various procedures will be illustrated using simulated series and real series produced by the Bureau of Labor Statistics in the U.S.A.

Controlled Hot- and Cold-Deck Imputation Algorithms

✦Yves TillÈ, University oh Neuchatel, Pierre ‡ Mazel 7, Neuchatel, 2000 Switzerland, *yves.tille@unine.ch*

Key Words: donor, imputation, sampling

Imputation methods are commonly and frequently used by survey practitioners. Our aim is to propose new, efficient, methods of imputation. We will focus on donor imputations that are a family of imputation methods where the missing value of a nonrespondent is replaced by the value observed for a respondent. Our method consists of formalizing imputation techniques as special sampling designs. Indeed, an imputation by donor can be viewed as the sampling of donors within the respondent population, where the order in the sample gives the correspondence between donors and receivers. We want to prepare for imputation by the construction of matrices of imputation that contain the probabilities for a nonrespondent to have his value imputed by that of a particular donor. Modern sampling techniques will then be applied on these matrices to realize an imputation. We establish a list of recommendations that enable us to evaluate the quality of an imputation method. Most of these recommendations can be written as constraints on these matrices of imputation and on the sampling designs that are used to make the imputation. Generalized calibration techniques and modern sampling techniques such as b

450 The Implications of Statistical Measurement on Public Policy ■●

Scientific and Public Affairs Advisory Committee, International Indian Statistical Association, Section on Government Statistics, Section on Health Policy Statistics, Statistics Without Borders

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

The Implications of Statistical Measurement on Public Policy

✦Amy Braverman, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, *amy.j.braverman@jpl.nasa.gov*; ✦Daniel F McCaffrey, RAND, 4570 Fifth Avenue, Suite 600, Pittsburgh, PA , *danielm@rand.org*; ✦James Rosenberger, Penn State University, University Park, PA 16802, *jlr@stat.psu.edu*; ✦Clyde Tucker, (Retired), 2 Massachusetts Ave., Washington, DC , *tucker.clyde@bls.gov*

Policymakers are becoming more and more dependent on statistical information to describe current conditions and make policy. Much of this information is a result of complex statistical measurement issues that are not at all obvious to the policymakers, yet alternative measurements could lead to dramatically different policies. This panel will present statistical measurement issues on four selected policy areas: Dan McCaffrey will discuss methodological issues in the Race To The Top such as modeling teacher contributions to student achievement given the complex realities of students' instructional inputs. He will also discuss the challenges to assigning student achievement to teachers' training programs. Clyde Tucker will discuss defining economic concepts (e.g. unemployment) through data collection, adjustment, estimation and analysis to the content of official publications of results by the Federal statistical agencies. Jim Rosenberger will discuss the measurement issues confronting the new Office of Financial Research created as part of the Wall Street reforms. Amy Braverman will describe measuring greenhouse gases from space used for monitoring climate change and treaty verification.

451 High Dimensional Data Analysis and Covariance Estimations ●

Section on Statistical Learning and Data Mining, International Indian Statistical Assoc., Reps. for Young Statisticians, Section on Statistical Computing

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Robust Functional Singular Value Decomposition Method

✦Lingsong Zhang, Purdue University, 150 N. University St, West Lafayette, IN 47907 United States, *lingsong@purdue.edu*

Key Words: functional data analysis, smoothing, robust method, outliers, GCV

Motivated by the analysis of a two-way functional data, we propose a novel robust functional singular value decomposition method. The regularized part of this method smoothes the estimated singular column and singular row. It also downweights the outlying effects slightly. The robustness part further reduces the outlying effects. A GCV method is developed for the smoothing parameter selection. Simulations are conducted to illustrate the usefulness of this new method.

A Path-Following Algorithm for Sparse Pseudo-Likelihood Inverse Covariance Estimation (SPLICE)

✦Guilherme Rocha , Indiana University, 309 N. Park Ave., Bloomington, IN 47408, *gvrocha@indiana.edu*; Peng Zhao, Citadel Investment Group; Bin Yu, UC Berkeley

Key Words: pseudo-likelihood, sparsity, regularization, covariance matrix

Given n observations of a p-dimensional random vector, the covariance matrix and its inverse (precision matrix) are needed in a wide range of applications. Sample covariance (e.g. its eigenstructure) can misbehave when p is comparable to the sample size n. Regularization is often used to mitigate the problem. In this paper, we proposed an l1-norm penalized pseudo-likelihood estimate for the inverse covariance matrix. This estimate is sparse due to the l1-norm penalty, and we term this method SPLICE. Its regularization path can be computed via an algorithm based on the homotopy/LARS-Lasso algorithm. Simulation studies are carried out for various inverse covariance structures for p=15 and n=20, 1000. We compare SPLICE with the l1-norm penalized likelihood estimate and a l1-norm penalized Cholesky decomposition based method. SPLICE gives the best overall performance in terms of three metrics on the precision matrix and ROC curve for model selection. Moreover, our simulation results demonstrate that the SPLICE estimates are positive-definite for most of the regularization path even though the restriction is not enforced.

Convergence and Prediction of Principal Component Scores in High-Dimensional and Ultra-High Dimensional Settings

✦Seunggeun Lee, Dept. of Biostatistics, Harvard University, Building 2, Room 451, 655 Huntington Avenue, Boston, MA 02115, *sglee@hsph.harvard.edu*

Key Words: PCA, PC scores, Random Matrix

A number of settings arise in which it is of interest to predict Principal Component (PC) scores for new observations using data from an initial sample. In this talk, we demonstrate that naive approaches to PC score prediction can be substantially biased towards 0 in the analysis of large matrices. This phenomenon is largely related to known inconsistency results for sample eigenvalues and eigenvectors as both dimensions (p) and sample sizes (n) increases. For the spiked eigenvalue model for random matrices, we expand the generality of these results, and propose bias-adjusted PC score prediction. Simulation and real data examples from the genetics literature show the improved bias and numerical properties of our estimators. In addition, we discuss asymptotic behaviors of sample eigenvalues, eigenvectors and PC scores under high dimensional (p/n < 8) and ultra high dimensional (p/n ? 8) settings.

Principal Subspace Estimation in Spiked Covariance Models

✦Zongming Ma, Department of Statistics, University of Pennsylvania, 400 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104, *zongming@wharton.upenn.edu*

Key Words: high dimension, principal component analysis, sparsity, spiked covariance model, subspace

For high-dimensional data, it is often desirable to reduce the dimensionality by projection onto a low dimensional principal subspace. However, classical PCA usually cannot find the subspace consistently in high dimensions. In this talk, we present a new principal subspace estimation method. For a class of spiked covariance models with sparsity constraints, it consistently, and even optimally, estimates the subspace.

452 Advances in High Dimensional Data Analysis and Applications in Biostatistics ●

Biometrics Section, ENAR, Section for Statistical Programmers and Analysts, Section on Statistics in Epidemiology

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

A Tale Of Two Manifolds

✦Sayan Mukherjee, Duke University, Durham, NC 27708 USA, *sayan@stat.duke.edu*

Key Words: Supervised dimension reduction, Manifold learning, Bayesian inference, Factor models

The focus is on the problem of supervised dimension reduction (SDR). We first formulate the problem with respect to the inference of a geometric property of the data, the gradient of the regression function with respect to the manifold that supports the marginal distribution. We provide an estimation algorithm, prove consistency, and explain why the gradient is salient for dimension reduction. We then reformulate SDR in a probabilistic framework and propose a Bayesian model, a mixture of inverse regressions. In this modeling framework the Grassman manifold plays a prominent role.

Functional Additive Regression

✦Yingying Fan, University of Southern California, Hoffman Hall, Information and Operations Management, Los Angeles, CA 90089 USA, *fanyingy@marshall.usc.edu*; Gareth James, University of Southern California

Key Words: Variable Selection, Functional Regression, Additive Model, Single Index Model, Shrinkage

We suggest a new method, called "Functional Additive Regression", or FAR, for efficiently performing high dimensional functional regression. FAR extends the usual linear regression model involving a functional predictor, X(t), and a scalar response, Y, in two key respects. First, FAR uses a penalized least squares optimization approach to efficiently deal with high dimensional problems involving a large number of different functional predictors. Second, FAR extends beyond the standard linear regression setting to fit general non-linear additive models. We demonstrate that FAR can be implemented with a wide range of penalty functions using a highly efficient coordinate descent algorithm. Theoretical results are developed which provide motivation for the FAR optimization criterion. Finally, we show through simulations and a real data set that FAR can significantly outperform competing methods.

Modified Covariate Approach For Personalized Medicine

✦Lu Tian, Stanford University, Redwood Building, Stanford University, Palo Alto, CA 94303, *lutian@stanford.edu*

Key Words: personalized medicine, dimension reduction, treatment-covariate interaction

It is important to stratify patient population such that the treatment effect is different for different subgroups of patients. The stratification can be used to select the optimal treatment for individual patients in a given subgroup. Statistically speaking, it amount to identify treatment and covariate interactions. This is a challenging task especially when the dimension of covariate is high. We propose a novel approach to directly estimate the covariate-treatment interaction without the need for modeling the main effect. Therefore, the method is very appealing for cases where the high-dimensional covariate such as gene expression levels is involved. We provide some theoretical justifications from casual inference perspective and illustrate the proposed with real data examples.

Testing For The Effect Of A Genetic Pathway In Longitudinal/Clustered Data Using Kernel Machine Regression

✦Arnab Maity, North Carolina State University, Department of Statistics, Campus Box 8203 , 2311 Stinson Drive, Raleigh, NC 27695, *amaity@ncsu.edu*; Stacey Alexeeff, Harvard University

Key Words: Longitudinal/clustered data, Kernel Machine regression, Score test, Gene-gene interaction

There is a growing scientific interest to test for genetic effects on disease by considering a set of genes that may be on the same biological pathway. Genes within a pathway may interact in functional ways to influence the progression of disease. Kernel machine regression has been introduced as a way to model a pathway effect, either parametrically or nonparametrically. We consider kernel machine regression for the testing of a genetic pathway effect in the longitudinal/clustered data setting, where a continuous disease outcome is measured repeatedly, possibly over time, for each subject. We develop a score-based test statistic for testing the effect of the genetic pathway accounting for the within subject correlation in the outcome variable. In addition, we present a simulation study to investigate the power of the test for different correlation structures, and we compare its performance with the test without accounting for correlation.

Population Functional Data Analysis Of Group Ica-Based Connectivity Measures From Fmri

✦Shanshan Li, Johns Hopkins University, MD 21205, *shli@jhsph.edu*; Brian Caffo, Johns Hopkins Department of Biostatistics; Suresh Joel, Johns Hopkins University; Stewart Mostofsky, Johns Hopkins University; Jim Pekar, Johns Hopkins University; Susan

Bassett, Johns Hopkins University

Key Words: Two-Stage Decomposition, ICA, PCA, fMRI

In this manuscript, we use a two-stage decomposition for the analysis of functional magnetic resonance imaging (fMRI). In the first stage, spatial independent component analysis is applied to the group fMRI data to obtain common brain networks (spatial maps) and subject-specific mixing matrices (time courses). In the second stage, functional principal component analysis is utilized to decompose the mixing matrices into population-level eigenvectors and subject-specific loadings. Inference is performed using conditional logistic regression for matched pairs data. Simulation studies suggest the ability of the decomposition methods to recover population brain networks and the major direction variation of the mixing matrices. The method is applied to a novel fMRI study of Alzheimer's disease risk under a verbal paired associates task. We found empirical evidence of alternative ICA-based metrics of connectivity in clinically asymptomatic at risk subjects when compared to controls.

453 Using the American Community Survey as a Sampling Frame - A Collaborative Effort ■

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Nscg Sampling Issues When Using An Acs-Based Sampling Frame

◆David Hall, U.S. Census Bureau, 14206 Oxford Dr., Laurel, MD 20707, *halldw@gmail.com*; Stephen Cohen, National Science Foundation; John Finamore, National Science Foundation; Flora Lan, National Science Foundation

Key Words: NSCG, ACS, Sampling Frame

The National Survey of College Graduates (NSCG) is a source of detailed statistics on the nation's science and engineering labor force. The Census Bureau conducts the NSCG on behalf of the National Science Foundation (NSF). Historically, the NSCG selected its sample once a decade from the decennial census long form respondents. In 2010, the Census discontinued the long form, so the NSF switched to using the American Community Survey (ACS) as a sampling frame for the NSCG. This switch brought both opportunities and challenges. The on-going data collection methods of the ACS offered the potential for improved coverage of the NSCG, and the inclusion of a new field of degree question on the ACS allowed for a more efficient sampling effort. Meanwhile, the ACS also introduced increased weight variation and complexity. This paper provides insight into our efforts to address the challenges of an ACS-based sampling frame by discussing the 2010 NSCG sample design issues, including the determination of an appropriate sample selection technique, the incorporation of an iterative sample allocation approach, and the identification and resolution of obstacles encountered during sample selection.

Nscg Estimation Issues When Using An Acs-Based Sampling Frame

◆John Finamore, National Science Foundation, 4201 Wilson Blvd, Arlington, VA 22230 United States, *jfinamor@nsf.gov*; Stephen Cohen, National Science Foundation; David Hall, U.S. Census Bureau; Donsig Jang, Mathematica Policy Research, Inc.; Julie Walker, U.S. Census Bureau

Key Words: NSCG, Multple Panel Estimation, ACS

The National Survey of College Graduates (NSCG) is the nation's only source of detailed statistics on the science and engineering labor force. Historically, the NSCG selected its sample once a decade from the decennial census long form respondents. In the 2010 NSCG survey cycle, the NSCG began using the American Community Survey (ACS) as the sampling frame for the NSCG. After considering numerous sample design options proposed by the NSCG survey sponsor, the National Science Foundation (NSF), and reviewed by the Committee on National Statistics (CNSTAT), the NSF approved the use of a rotating panel design for the 2010 decade of the NSCG. This rotating panel design allows the NSCG to address certain deficiencies of the previous long form-based design including the undercoverage of key interest groups. However, along with numerous improvements, the use of the ACS as a sampling frame for the NSCG and the implementation of the NSCG rotating panel design also introduced new challenges. This document summarizes the rotating panel design planned for the 2010 decade of the NSCG and discusses results from two research tasks related to NSCG estimation.

Evaluating 2003 Nscg Dual Frame Estimates For 2010 Nscg Planning Purposes

◆Donsig Jang, Mathematica Policy Research, Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, *djang@mathematica-mpr.com*; David Hall, U.S. Census Bureau

Key Words: NSCG, Dual Frame Estimation

The National Survey of College Graduates (NSCG) is the nation's leading source of detailed statistics on the science and engineering labor force. Beginning in the 2010 survey cycle, the NSCG will be constructed using multiple sampling frames. The NSCG had attempted a similar dual frame approach in the 2003 NSCG survey cycle, but differing population estimates between the frames led the survey sponsor, the National Science Foundation (NSF), to abandon the dual frame estimates for publication purposes in favor of the single frame estimates. New research into the 2003 NSCG dual frame design has presented an opportunity to reevaluate the 2003 estimates and the 2003 decision. This paper looks deeper into the 2003 NSCG dual frame design issues, including potential causes of the differing estimates and comparisons of single frame and dual estimates two frames for key estimates of interest. The goal of this research is to better understand the 2003 dual frame estimates in preparation for the 2010 NSCG.

Nscg Variance Estimation - Variance Estimation In A Two-Phase Sample Design Setting

◆Michael Edmund White, U.S. Census Bureau, 4700 Silver Hill Rd, Suitland, MD 20746, *michael.e.white@census.gov*; Jean Opsomer, Colorado State University

Key Words: NSCG, ACS, two-phase sample design, variance estimation

The National Survey of College Graduates (NSCG) is a longitudinal survey that collects information on employment, educational, and demographic characteristics of the college-educated science and engineering workforce in the United States. The U.S. Census Bureau conducts the NSCG on behalf of the National Science Foundation (NSF). Historically, the NSCG sample was selected once a decade from the decennial census long form respondents. In 2010, the Census discontinued the long form, so the NSF switched to using the American Community Survey (ACS) as a sampling frame for the NSCG. The two-phase sample design of the NSCG creates variance estimation complexities. This paper will explore the variance estimation issues we faced in using the ACS as a sampling frame for the NSCG. As part of the research, simulation studies compared variance estimators under numerous replication methods including successive difference, grouped-jackknife, and balanced repeated replication in an effort to determine which replication method best met the NSCG estimation needs.

454 Statistical Use of Administrative Records from non-Federal Sources ■●

Section on Government Statistics, Social Statistics Section, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Statistical Use Of Administrative Records From Non-Federal Sources

◆Anne Stahl, SocialSecurity Administration, 500 D Street, SW, 9th Floor, Washington, DC 20254 US, *anne.stahl@ssa.gov*

Key Words: Supplemental, Security, Income, Temporary, Assistance, Families

The TANF-SSI Disability Transition Project (TSDTP) is a research partnership between HHS's Administration for Children and Families (ACF), the Social Security Administration (SSA), and several state and local authorities, to better coordinate the Temporary Assistance to Needy Families (TANF) and Supplemental Security Income (SSI) programs. One goal of the project includes analysis of TANF and SSA administrative data to understand the overlap of the two program populations, SSI application patterns, and eligibility and employment outcomes. This project, initially planned in 2006, was motivated in part by GAO's TANF and SSI: Opportunities Exist to Help People with Impairments Become More Self-Sufficient (2004) report recommending that SSA begin a demonstration project to improve services and coordination with TANF agencies. Gaining access to and analyzing state TANF administrative data has proven very challenging due to a variety of legal, bureaucratic and technical issues. The project has also gained access to ACF's national TANF data. These data are being linked to SSA disability data to get a broader representation of the TANF-SSI interaction.

Building A Comprehensive Data Set On Every Teacher: The Nces Teacher Compensation Survey

◆Stephen Q. Cornman, US Dept. of Education, NCES, 1990 K Street NW, Room 9104, Washington, DC, DC 20006, *stephen.cornman@ed.gov*; Frank Johnson, US Dept. of Education, NCES; Lei Zhou, Educational Services Institute-MacroSys, LLC; Amber Noel, American Institutes for Research

Key Words: Administrative data on teachers' compensation, teacher salaries, teacher characteristics

National data on teachers are limited to periodic sample surveys or to simple counts at the district or school level. In response to the need for individual teacher-level data, the US Dept. of Education, NCES developed the Teacher Compensation Survey (TCS), an administrative records survey that collects total compensation, teacher status, and demographic data about individual teachers from multiple states. Approximately 1.6 million teachers are currently in the data set, representing 50% of teachers in the United States. In 2007, NCES launched the pilot TCS data collection, with seven states volunteering to provide administrative records for school year (SY) 2005-06. The TCS expanded to 17 states reporting SY 2006-07 data, 18 states reporting SY 2007-08 data, and 23 states reporting SY 2008-09 data. It is anticipated up to thirty-five states will volunteer to participate in the TCS from 2011 to 2013. This session provides an overview of the TCS data collection, a comparison of state administrative records with other sources of data, data availability and quality, limitations, and advantages of the TCS. This session also presents findings and descriptive statistics.

Higher Education Effects On Employment Of Deaf

◆Robert Weathers, Social Security Administration, Office of Program Development and Research, SSA, 6401 Security Boulevard, 128 Altmeyer, Baltimore, MD 21235 U.S.A., *Robert.Weathers@ssa.gov*

We use administrative data from the Rochester Institute of Technology (RIT), the National Technical Institute for the Deaf (NTID), the Social Security Administration (SSA), and the Rehabilitation Services Administration (RSA) to investigate the degree to which federal government investments in higher education for deaf students (via Education of the Deaf Act and the federal/state VR program) will reduce their use of Supplemental Security Income (SSI)/Social Security Disability Insurance (SSDI) by increasing their long term employment and earnings. Specifically, this project will test the hypotheses that federal investments in education increased the long term earnings of deaf students who attended the National Technical Institute for the Deaf (NTID) at Rochester Institute of Technology (RIT) and 1) compare their long term employment and earnings with a matched sample of RIT hearing students, 2) compare their long term employment and earnings with a sample of deaf students who received federal/state VR funding for higher education, and 3) show how the Great Recession of 2008-2010 affected the employment and earnings of these three populations.

Summary Of Current Studies On Homeless Populations Within The Social Security Administration Databases

◆Robert Pfaff, Social Security Administration, Ste. 122 Altmeyer, 6401 Security Blvd., Woodlawn, MD 21235, *Robert.Pfaff@ssa.gov*

Key Words: Homeless, Statistical, Analysis

To present project summaries and provide information regarding research opportunities utilizing transient data from the Social Security Administration's administrative databases. The first project, the Homeless Outreach Project Evaluation (HOPE) Longitudinal Study, has completed the outreach and evaluation utilizing a retrospective cohort of transients located in the administrative databases. The second project, the San Francisco and Santa Cruz Counties Presumptive Disability Pilot Project, will test the viability of offering Presumptive Disability (PD) determinations to the homeless population with a diagnosis of schizophrenia. The presumptive disability process would allow homeless individuals with schizophrenia to start receiving SSI benefits immediately, prior to a full medical review by the state Disability Determination Services (DDS). The third project, the LA County Benefits Entitlement Services Team (BEST) Project Evaluation will study homeless participants receiving local health care and outreach services through four locations in Los Angeles County. In an effort to improve the overall conditions of these homeless individuals, the County of Los Angeles ap

455 Monitoring and Change Detection in Industrial/Health Quality Control and Related Topics ■●

Section on Quality and Productivity

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Sensitivity Analysis Of Predictive Modeling For Responses From The Three-Parameter Weibull Model With A Follow-Up Doubly Censored Sample Of Cancer Patients

◆Saeed Alshahrani, Florida International University, Department of Epidemiology & Biostatistics, Robert Stempel College of Public Health, Miami, FL 33199 USA, *salsh001@fiu.edu*; Ahmed Albatineh, Florida International University ; Nasar U. Ahmed, Florida International University ; Nadine Jenkins, University of Medicine & Dentistry of New Jersey; Hafiz M. R. Khan, Florida International University

Key Words: Doubly censored sample, Highest predictive density, Goodness of fit tests, Three-parameter Weibull model, Bayesian approach, Predictive inference

The purpose of this paper is to derive the predictive densities for future responses from a three-parameter Weibull model given a doubly censored sample. The predictive density for a single future response, bivariate future response, and a set of future responses have been derived when the shape parameter a is known and unknown. A real data example representing 44 patients who were diagnosed with laryngeal cancer (2000-2007) at a local hospital is used to illustrate the predictive results for the four stages of cancer. The survival days of eight out of the 44 patients could not be calculated as the patients were lost to follow-up. These were the first four and the last four patients. Thus, the recorded data for the survival days of 36 patients composed of 18 male and 18 female patients with cancer of the larynx are used for the predictive analysis. Furthermore, a subgroup level of the male and female patients follow-up data are considered to obtain the future survival days. A sensitivity study of the mean, standard deviation, and 95% highest predictive density interval of the future survival days with respect to stages and doses are performed when the shape parameter a is unknown.

Monitoring Process Capability Indices

◆Ejaz Syed Ahmed, University of Windsor, windsor, ON n9b 3p4 canada, *seahmed@uwindsor.ca*; Abdul K. Hussein, University of Windsor

Key Words: Process Capability, Sequential testing, Brownian motion

A Process Capability Index (PCI) is a numeric summary that compares the behavior of a product or process characteristics to engineering specifications. We propose a sequential procedure for testing whether two processes are equally capable by using. We employ a nonsequential Wald-type statistic and provide its sequential version by Brownian motion approximations. We point out that, as a byproduct, the nonsequential Wald-type statistic used here, provides an easily computable alternative to Boyels' approximate F-test (Boyels 1991). We give an algorithm for conducting the sequential test and we examine its performance by Monte Carlo simulations. Finally, we illustrate the method by testing capability improvement of an industrial process before and after calibration based on published data.

Nonparametric Sequential Monitoring Of Longitudinal Trials

◆Edit Gombay, University of Alberta, Edmonton, AB Canada, *egombay@ualberta.ca*

Key Words: Sequential Monitoring, Longitudinal trials, Brownian Motion, Group sequential

We considers the sequential monitoring of multi-armed longitudinal clinical trials. We describe an approach that is relatively simple and accessible. Sequential ranks are used to form partial sum statistics, yielding processes that have independent increments, and hence can be approximated by Brownian motions. Three monitoring procedures are proposed. The first two are asymptotic, continuous analogues of the well-known Pocock and O'Brien-Fleming group sequential procedures, while the third procedure is exact. Performance of the procedures is assessed using Monte Carlo simulations. Data from an orthodontic clinical trial is used to illustrate the proposed methods, for the comparison of three treatment groups.

Effect Of Correlations On Risk-Adjusted Cusum Charts For Monitoring Binary Outcomes

◆Abdul K. Hussein, University of Windsor, Windsor, ON Canada, *ahussein@uwindsor.ca*; Abdullah Kasem, University of Windsor; Severien Nkurunziza, University of Windsor; Stefano Campostrini, Ca Foscari University

Key Words: CUSUM, Autocorrelation, Binary outcomes

Recently, monitoring health care outcomes, such as post-operative mortality rates has become common and new statistical methodologies have started to emerge for this purpose. One of such methodologies is the risk adjusted CUSUM chart (RA-CUSUM) for monitoring binary outcomes such as mortality after cardiac surgery. However, it is common to assume that the observations being monitored, which often come from the same surgeon or institution, are independent. In this talk, we examine the effect of various types of autocorrelations on the performance of the RA-CUSUM charts in terms of their Average Run Lengths (ARL). We give a brief discussion of other possible shortcomings of the RA-CUSUM and remedies for these issues.

New Cusum Charts For Monitoring Binomial Outcomes

✦Severien Nkurunziza, University of Windsor, Windsor, ON n9b 3p4 Canada, *severien@uwindsor.ca*; Shabnam Chitsaz, University of Windsor; Saber Fallahpour, University of Windsor; Abdul K. Hussein, University of Windsor; Saber Fallahpour, University of Windsor

Key Words: Binomial, Poisson, CUSUM, non-conforming, C-section

Classically, it has been common in industry to employ CUSUM charts for monitoring non-conforming fractions which follow binomial distribution. Here we extend this methodology to accommodate binomial distributions in which the number of trials are variable. We employ Binomial-Poisson compound distribution and develop CUSUM chart threshold tables for desired ARL under both in-control and out-of-control processes. We apply the method to data on C-section deliveries.

456 Disclosure Avoidance for Tabular Data ■●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section, Committee on Privacy and Confidentiality, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Initial Efforts To Extend The Ezs Noise Methodology To The Quarterly Census Of Employment And Wages

✦Michael Buso, Bureau of Labor Statistics, 2 Massachusetts Avenue, NE Room 4840, Washington, DC 20212, *buso_m@bls.gov*

Key Words: tabular data, disclosure, noise, random noise

The Bureau of Labor Statistics (BLS) Quarterly Census of Employment and Wages (QCEW) program has been using primary and complementary cell suppression techniques as its disclosure limitation methodology. The employment and wage levels of over 9 million establishments are collected under confidentiality pledges and other provisions that obligate BLS to protect the identities and characteristics of over 9 million establishments whose data are assembled into a multi-dimensional array of over 3.5 million populated data cells. The techniques have been providing an unsatisfactory balance between the protection provided to respondents and the usefulness of the data for economic and policy analysis uses, as over 2.1 million cells are suppressed. BLS is seeking to extend the basic random noise methods of Evans, Zayatz, and Slanta (EZS), to the characteristics of the QCEW program. BLS also is considering whether several of the protective mechanisms of the original EZS paper can be relaxed while still providing reasonable protection to respondents.

Applying Cell Suppression to Inter-Related Tables

Jacob Bournazian, U.S. Energy Information Administration; Michael Salpeter, Addx Corporation; ✦Bin Zhang, Energy Information Administration, 1000 Independence Ave SW EI-21, Washington DC, DC 20585, *bin.zhang@eia.gov*

Key Words: confidentiality, cell suppression

Cell suppression is the most common disclosure limitation method that the U.S. Energy Information Administration (EIA) applies to the aggregate statistics that it publicly releases. Complementary cell suppression involves withholding the publication of non-sensitive cells in a table in order to protect the cells that were identified as sensitive to revealing company level information. EIA has two different automated suppression programs that it applies to tabular data. Each program follows a different methodology. The data protection levels for a table vary from applying different complementary cell suppression methodologies because the suppression patterns vary depending upon each methodology. If several tables within the same information product are related due to a high level of dimensions in the table designs, then the selection of complementary cells in each table becomes a more complicated computational problem. This paper focuses on the situation when tables are related and compares two different automated approaches for applying cell suppression methodology.

Modernizing Cell Suppression Software At The U.S. Census Bureau

✦Paul B. Massell, U.S. Census Bureau, 4700 Silver Hill Road, Room 5K114A; CDAR, Washington, DC 20233, *paul.b.massell@census.gov*

Key Words: Cell Suppression, Sensitive Cells, Protection at the Company Level

An Economic Census of the United States is conducted every five years by the economic directorate of the U.S. Census Bureau. The main data products are additive magnitude data tables that typically involve NAICS categories as rows and geographic entities as columns (some tables have a 3rd dimension). The p% rule is used for determining which cells are sensitive, and how much protection each such cell requires. First, the sensitive cells are suppressed. Then a cell suppression program is run against a file with information about each cell including an identifier for each establishment that contributes to the cell value; its associated company, and the contributed value. This program calls an optimization routine for each sensitive cell in order to find the optimal set of additional cells that must be suppressed in order to find the set of cells with the minimum total value, that, when suppressed, lead to a protection of the sensitive cells at the company level. We discuss a number of complex aspects of the software and how each of these was modernized.

Using Tau-Argus And R-Statistical Package Sdctable To Conduct Secondary Cell Suppression For Linked Tables

✦Amang Sukasih, Mathematica Policy Research, 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, *asukasih@mathematica-mpr.com*; Donsig Jang, Mathematica Policy Research, Inc.; David Edson, Mathematica Policy Research

Key Words: disclosure avoidance, tabular data

When a data cell in a table is suppressed by dropping its value based on a primary cell suppression rule, the value of that cell can still be determined if the table, subtable, or linked tables provide totals, marginal totals, or subtotals. Secondary cell suppression is therefore needed to avoid such disclosures. Two software packages are available to assist researchers with secondary cell suppression: Tau-Argus (Statistics Netherland 2009) and R-statistical package sdcTable (Meindl 2010). But even with this software, there is no simple way to perform secondary suppression for linked tables-that is, tables presenting data on the same cells that share some categories of at least one explanatory variable. Computation may not be trivial and may still require manual reviews, especially when dealing with a large number of linked tables. With an eye toward finding the most straightforward and effective method of suppressing linked tables, we will explore the capabilities of the aforementioned software programs in performing linked-table suppression, identifying the strengths and limitations of each program and comparing the results.

457 Statistics in the Environmental and Geophysical Sciences ●

Section on Statistics and the Environment, Section on Physical and Engineering Sciences, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Gridded Forecast Verification

✦Eric Gilleland, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, *ericg@ucar.edu*

Key Words: forecast verification, spatial statistics, image analysis, high-resolution weather forecast modeling, image warping, band-pass filtering

As numerical weather forecasts continue to increase their resolution, they have become more informative. To most subjective users, they are superior. However, traditional forecast verification, which aggregates grid-point to grid-point comparisons over an entire grid, often favor the coarser (smoother) scale models. Subsequently, many new approaches have been proposed, many of which have been inter-compared through the Spatial Forecast Verification Methods Inter-Comparison Project (ICP, http://www.rap.ucar.edu/projects/icp). This talk will give an overview of the major new approaches, with emphasis on a technique originating in image analysis, image warping. Further, it will be shown how this technique can potentially be advanced to incorporate space and time in this setting.

Spatio-Temporal Exceedance Locations And Confidence Regions

✦Joshua P French, University of Colorado Denver, Campus Box 170, P.O. Box 173364, Denver, CO 80217, *joshua.french@ucdenver.edu*; Stephan Sain, NCAR

Key Words: exceedance regions, hotspots, confidence regions, environmental monitoring, spatio-temporal

An exceedance region is the set of locations in a spatial domain where a process exceeds some threshold. Examples include areas where ozone concentrations exceed safety standards; there is high risk for tornadoes or floods; or heavy-metal levels are dangerously high. Identifying these regions in a spatial or spatio-temporal setting is an important responsibility in environmental monitoring. Exceedance regions are often estimated by finding the areas where an estimated model exceeds some threshold. Even when estimation error is quantifiable at individual locations, the overall estimation error of the estimated exceedance region is still unknown. A method is presented for constructing a confidence region containing the true exceedance region of a spatio-temporal process at a future time. Conventional techniques are used to model the spatio-temporal data, and then conditional simulation is combined with hypothesis testing to create the desired confidence region. A simulation study is used to validate the approach for several levels of spatial and temporal dependence. The methodology is applied in assessing climate change using data from the National Center for Atmospheric Research.

A Gaussian Markov Random Fields Approach To Integrating Multiple Climate Model Output

✦E. Shannon Neeley, Brigham Young University, 223 TMCB, Provo, UT 84602, *sneeley@stat.byu.edu*; William F. Christensen, Brigham Young University; Stephan Sain, NCAR

Key Words: Gaussian Markov Random Fields, Regional Climate Models, NARCCAP, Factor Analysis, Spatial Prediction

Multi-model ensembles are commonly used in climate research to help quantify and understand single model uncertainties. There have been various published methods for creating these ensembles often by taking simple or weighted averages within a Bayesian hierarchical model framework. Christensen and Sain (Math Geosci, 2011) consider an alternative approach to multi-model ensembles using a spatially correlated latent variable model. Rather than providing a weighted average, their approach highlights similarities among model ensemble members and provides feedback about which models differ most from the group and why. The current work extends and improves their method with a model that uses a single smoothed factor analysis rather than running an independent factor analysis at each site and smoothing afterward with heterogeneous variance measurement-error-filtered kriging. We use Gaussian Markov Random Fields (GMRF) in conjunction with spatial factor analysis. We apply our method to the North American Regional Climate Change Assessment Program (NARCCAP). Our model still allows assessment of model similarities while incorporating a more realistic smoothing method.

Combining Multiple Computer Models For Uncertainty Quantification In Posterior Predictive Inference

✦Matthew T. Pratola, Los Alamos National Laboratory, , *mpratola@lanl.gov*

Key Words: computer experiments, bayesian model averaging, climate models, latent variable, uncertainty quantification

Computer models enable scientists to investigate real-world phenomena in a virtual laboratory using computer experiments. Frequently, multiple computer models are available which make different assumptions in attempting to simulate the physical process of interest. Combining multiple computer models has not been well-addressed in the computer experiments literature. We develop a Bayesian multi-model statistical framework to address this problem. In particular, our approach describes a latent model-space, which may be integrated out to fully incorporate model uncertainty in the posterior inference.

Calibrating Computer Models With Multivariate High And Low Resolution Output

✦William Kleiber, National Center for Atmospheric Research, CO , *wkleiber@ucar.edu*; Stephan Sain, NCAR

Key Words: Computer experiment, Calibration, Fidelity, Resolution, Spatial-temporal

We consider calibration of the Lyon-Fedder-Mobarry model for coupled magnetosphere-ionosphere simulation. Model output is a bivariate spatial-temporal field for two physical variables, and is available at high and low resolutions. We discuss an approach to quantifying the uncertainty in the multivariate calibration parameters that combines the computationally expensive but sparser high resolution output with the lower fidelity but computationally inexpensive low resolution output.

458 Statistical Considerations for the Evaluation of Diagnostic Devices and Biomarkers ■

Biopharmaceutical Section, Section on Health Policy Statistics, Section on Risk Analysis, Section on Statistics in Epidemiology, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Early Detection Claim In Ivd Devices: Statistical Issues And Challenges

✦Chunrong Cheng, FDA, R471S 1401 Rockville Pike, Rockville, MD 20852 USA, *chunrong.cheng@fda.hhs.gov*; Tie-Hua Ng, FDA

Key Words: IVD, early detection, window period

The "window period" is the time of infection of a virus (e.g., HIV) to the time that the infection could be detected by an in vitro diagnostics (IVD) devices. Shortening the window period could mean early detection. Claim for an early detection can be made by improving the IVD devices and technologies amplifying antigen/antibody production. However, assessing early detection faces major challenges such as study design, sample selection, sample size, etc. We will discuss the related statistical issues from a regulatory perspective. We hope to stimulate a discussion that may lead to solutions to these challenges.

Prediction Biomarkers For Safety Or Efficacy

✦Estelle Russek-Cohen, US FDA Center for Biologics, 1401 Rockville Pike HFM-215, Rockville, MD 20852, *Estelle.Russek-Cohen@fda.hhs.gov*; Rong Tang , US FDA Center for Devices

Key Words: Prediction Biomarker, safety, efficacy, study design

Most of the statistical literature regarding evaluation of prediction biomarkers that identify patients for a particular therapy focus on efficacy of the therapy. However, we now see markers (such as those assessed using in vitro assays) that are used to identify patients at greater risk for an adverse event associated with the use of the therapy. The study designs and endpoints for these kinds of studies are apt to be different from those evaluating efficacy and may also vary depending on whether the therapy is already in use. We will present an overview of some of the design and analysis considerations, contrasting study designs for prediction biomarkers used for evaluating efficacy versus those used for addressing safety of a therapeutic.

Individualized Disease Diagnostics: Improving Diagnostic Tests With Genomic/Genetic/Clinical Information

✦Meijuan Li, FDA/CDRH, #2225/66, 10903 New Hampshire Avenue, Silver Spring, MD 20993 USA, *meijuan.li@fda.hhs.gov*

Key Words: Individualized Disease Diagnostics, cutoff, diagnostics medicine, predictive value, diagnostic accuracy, overdiagnosis and overtreatment

A patient's certain clinical or genetic/genomic characteristics may affect his diagnostic test level independent of his disease status. However, a uniform cutoff (i.e. a population based cutoff) regardless of patients' characteristics is often used in diagnostics medicine where a cutoff is defined as the threshold at which the test differentiates a positive from a negative outcome. Consequently, such test can be controversial because of its modest predictive value and diagnostic accuracy; the potential overdiagnosis and overtreatment of the disease. Therefore, using more individualized thresholds based on a test; molecular measures; and patient age, ethnicity, and family history of disease can improve the predictive ability and the accuracy of this test. In this talk, we will discuss issues related to population-based diagnostic approaches and discuss some statistical methods for individualized diagnostic approaches.

Minimum Statistical Standards For Biomarker Development

✦Jae Kyun Lee, University of Virginia, POBOX 800717, Charlottesville, VA 22911, *jaeklee@virginia.edu*

Key Words: biomarker, SBIG, StatBDS

Development of predictive molecular and clinicopathological biomarkers for disease diagnosis, prognosis and therapeutic response is critical for continuous efforts in personalized medicine. Recent biomarker developments have become extremely challenging statistically, especially since a vast amount of molecular and clinicopathological data are commonly utilized in these developments. The so-called Statistical Biomarker Interest Group (SBIG), consisting of >40 statistical biomarker scientists from FDA, industry, and academia, was formed in 2010. SBIG is currently making efforts to derive a consensus opinion on the most essential statistical issues and criteria for biomarker development, termed as "Minimum Statistical Biomarker Development Standards (StatBDS)."

459 Bayesian Theory and Methods

Section on Bayesian Statistical Science

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Student Paper Competitions

✦Hugo Andres Gutierrez, Universidad Santo Tomas, Av caracas No 49 55 apt 216, Bogot·, International Colombia, *psirusteam@gmail.com*

Key Words: Bayesian analysis, generalized linear models, heteroscedasticity, Markov chain Monte Carlo, mixed effects models, nonlinear regression

This paper is aimed at the development of a Bayesian approach to handle inference and estimation in a large class of statistical models, which unifies and extends many classical methods that have received a great deal of attention across a wide range of scientific disciplines. The class of generalized nonlinear models includes some of the main techniques of the statistical methodology, namely generalized linear models and parametric nonlinear regression. This paper develops a systematic methodology for inference and computation in such models. In addition, this proposal extends to methods for the systematic treatment of variation that is not explicitly predicted within the model, trough the inclusion of random effects, and takes into account the modeling of dispersion parameters when the two-parameter exponential family is used as baseline for the likelihood of data. The methodology is based in the implementation of a two-stage algorithm that induces an hybrid approach based on numerical methods for approximating the likelihood to a normal density and further Taylor linearization around the values of current parameters in an MCMC routine.

Bayesian Predictive Inference for a Matrix-T Errors Regression Model

✦Azizur Rahman, University of Canberra, NATSEM, University of Canberra, Canberra, International ACT 2601 Australia, *Azizur.Rahman@canberra.edu.au*

Key Words: Bayesian method, matrix-T distribution, multivariate simple regression, prediction distribution, tolerance region

Prediction distribution is a basis for predictive inference applied in many real world situations. The Bayesian method is utilized in this research to derive the prediction distribution of future responses matrix for multivariate linear model with matrix-T error. Results reveal that the prediction distribution of a future responses matrix is a matrix-T distribution with appropriate location, scale and shape parameters. The prediction distribution depends on the realized responses only through the sample regression matrix and the sample residual sum of squares and products matrix. The shape parameter of the prediction distribution depends on the size of the sample and the number of regression parameters of the model, but it does not depend on the shape parameter of the error distribution. It is noticeable that the study model is robust and the objective Bayesian method is competitive with other statistical methods in the derivation of prediction distribution. A Beta-expectation tolerance region for the future responses matrix of the regression model has been obtained. Some applications of predictive inference have also been illustrated.

Student Paper Competitions

✦Raphael Nishimura, Survey Methodology Program, University of Michigan, 426 Thompson Street, Room 4050, Ann Arbor, MI 48104, *raphaeln@umich.edu*; Rafael Bassi Stern, Department of Statistics, Carnegie Mellon University; Fernando Vieira Bonassi, Department of Statistical Science - Duke Univeristy

Key Words: Bayesian Statistics, Decision Theory, Game Theory, Intersubjectivism, Randomization

In research situations usually approached by Decision Theory, it is only considered one researcher who collects a sample and makes a decision based on it. It can be shown that randomization of the sample does not improve the utility of the obtained results. Nevertheless, we present situations in which this approach is not satisfactory. First, we present a case in which randomization can be an important tool in order to achieve agreement between people with different opinions. Next, we present another situation in which there are two agents: the researcher - a person who collects the sample; and the decision-maker - a person who makes decisions based on the sample collected. We show that problems emerge when the decision-maker allows the researcher to arbitrarily choose a sample. We also show that the decision-maker maximizes his expected utility requiring that the sample is collected randomly.

Bayesian Latent Variable Methods With Some Applications

✦Saman Muthukumarana, University of Manitoba, Department of Statistics, 338 Machray Hall, Winnipeg, MB R3T 2N2 Canada, *muthukum@cc.umanitoba.ca*; Tim Swartz, simon fraser university

Key Words: Latent variables, Goodness-of-fit, MCMC

In Bayesian statistics we are interested in the posterior distribution of parameters. In simple cases we can derive analytical expressions for the posterior. However in most situations, the posterior expectations cannot be calculated analytically in the existence of latent variables, multidimensional parameters, and complex model structures. When the resulting posterior distribution is complex and high-dimensional, posterior expectations can be approximated by MCMC methods. In this talk, I will present how Bayesian latent variable methods can be developed in diverse problems in the fields of sports, network and ordinal survey data. The proposed methodologies are tested through simulation studies. Goodness-of-fit procedures are also discussed for assessing the validity of the methods. In our applications, the data are high dimensional, and we advocate a comparison of prior-predictive output versus observed data via "features" using Euclidean distances. Goodness-of-fit procedures based on pivotal quantities are also discussed. The proposed goodness-of-fit procedures are intuitive and flexible in the sense that one can investigate features which are relevant to the particular model.

460 Improving the Ethical Guidelines

Committee on Professional Ethics

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Improving The Ethical Guidelines

✦John S. Gardenier, Independent Science/Statistical Ethicist, 115 Saint Andrews Drive NE, Vienna, VA 22180, *drgarden@verizon.net*; ✦Ron Wasserstein, Executive Director, American Statistical Association, 732 N.Washington Street,, Alexandria, VA 22314 USA, *ron@amstat.org*; ✦Peter B. Imrey, Cleveland Clinic and Case Western Reserve University, Dept. of Quantitative Health Sciences/JJN3, Cleveland Clinic, 8500 Euclid Ave., Cleveland, OH 44195, *imreyp@ccf.org*; ✦Douglas A Samuelson, InfoLogix, Inc. (Current Member, ASA Committee on Professional Ethics & Founding Member, ASA CSFHR), InfoLogix, Inc., 8711 Chippendale Court, Annandale, VA 2203 USA, *samuelsondoug@yahoo.com*

Key Words: Statistical Ethics, Ethical Guidelines, ASA Members

The 1999 Ethical Guidelines for Statistical Practice document was intended to receive a major rewrite every ten years or so. The revision process is now getting underway. Although the Committee on Professional Ethics was then and is now responsible for producing the final document, the contents reflect extensive input from the ASA members in general. Notably, about 100 people attended a Panel Discussion of the Guidelines at the 1998 JSM in Dallas. That meeting was crucially important in determining the contents that emerged one year later. This proposed panel has similar intent. Panelists will describe and discuss their preliminary thoughts on where the Guidelines might go next. They will also seek input from the audience on their experiences with statistical ethics, their reactions to the panelists' comments, and their recommendations for the scope and the conceptual content of the next guidelines document. Those inputs will be invaluable in helping the Committee to produce a updated document that suitably reflects the combined will and knowledge of the ASA membership.

461 Personalized Medicine and Convergence: Prospect for Statisticians ●

Section on Health Policy Statistics, ENAR

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Personalized Medicine and Convergence: Prospects for Statisticians

✦Turkan Kumbaraci Gardenier, Pragmatica Corp., 115 Saint Andrews Drive NE, Vienna, VA 22180, *gardeniert@yahoo.com*; ✦Ashley L Shade, Yale University, Kline Biology Tower, Room 308, 219 Prospect Street, New Haven, 06520, *ashley.shade@yale.edu*; ✦Sholom Wacholder, National Cancer Institute, Executive Plaza South, Room 8046Bethesda, MD 80892, *wacholds@mail.nih.gov*; ✦Knut Wittkowski, The Rockefeller University, 1230 York Ave Box 322, New York, NY 10065, *kmw@rockefeller.edu*; ✦Alan Shuldiner, University of Maryland School of Medicine, 666 West Redwood Street, Room 494, Baltimore, MD 21202, *ASHULDIN@medicine.umaryland.edu*

Key Words: genomics; microarrays; personalized medicine; convergence

Advances in genomics research has spurred many avenues for contribution by statisticians. Methodology for dealing with microarrays, false positives and false negatives in case control studies are some issues where statisticians interact with researchers in design and/or analysis and interpretation. "Convergence" as a related concept has also evolved, directing not only collaboration among disciplines but also has led to synergistic outcomes. Biologists, physical scientists and engineers are now working together in special institutes or university centers toward developing new products which can identify prospective onset of disease and biological interactions. Nanotechnology is often at the core of such areas of pursuit. This panel brings together contributors to the area of personalized medicine with a view to discussing and identifying newly emerging roles for statisticians-- for example, clinical trials for emerging products and related quality control issues, and on-site training in statistical data. analysis and data collection.

462 Preparing Students for the Future of Analytics ■●

Section on Statistical Education

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Preparing Students For The Future Of Analytics

✦Aric LaBarr, Institute for Advanced Analytics at North Carolina State University, 920 Main Campus Drive, Venture 2 Building, Suite 530, Raleigh, NC 27606 USA, *aric_labarr@ncsu.edu*; ✦Russell Zaretzki, The University of Tennessee Knoxville, , *rzaretzk@utk.edu*; ✦Jerry Oglesby, SAS, , *Jerry.Oglesby@sas.com*; ✦Dan Thorpe, Sam's Club, , *Daniel.thorpe@samsclub.com*

Key Words: Analytics, Statistical Education, Graduate Programs, Industry, Training

With the enormous amount of data being collected in this modern age, businesses are starting to adapt their decision-making processes to becoming more data driven. With this changing paradigm, businesses are expecting different abilities for students seeking careers as analysts. What are businesses looking for in these new students? Is academia sufficiently preparing modern analysts? What programs exist to help students prepare for this new environment? The panelists will address these concerns and develop new ideas for future solutions.

463 Issues with Missing Data in Longitudinal Studies

Biometrics Section, Biopharmaceutical Section, ENAR, International Indian Statistical Association, Section on Health Policy Statistics, Section on Statistics in Epidemiology, Section for Statistical Programmers and Analysts, Section on Survey Research Methods, Social Statistics Section, WNAR

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Analysis Of Longitudinal Binary Data With Nonignorable Dropout Using Shared Parameter Models

✦Myungok Lee, Louisiana State University Health Sciences Center, 1615 Poydras Street, New Orleans, LA 70112, *mlee1@lsuhsc.edu*; Keunbaik Lee, Louisiana State University Health Sciences Center

Key Words: Generalized linear models, Marginalized transition, Shared parameter models, Fisher-scoring

In longitudinal studies investigators frequently have to assess and address potential biases introduced by missing data. This paper proposes new methods for modeling longitudinal binary data with nonignorable dropout using maginalized transition models and shared parameter models. Random effects are introduced for both serial dependence of outcomes and nonignorable missingness. Fisher-scoring and Quasi-Newton algorithms are developed for parameter estimation. Methods are illustrated with a real data set.

Analysis Of Longitudinal Data With Non-Random Missingness Using Shared Random Effects Models

✦Xiaoyun Li, Merck & Co. Inc., 131, Church Road, Apt# 1J, North Wales, PA 19454 United States, *xiaoyun_li@merck.com*; STUART Lipsitz, Division of General Medicine, Brigham and Women's Hospital, Boston, MA; DIPANKAR BANDYOPADHYAY, Division of Biostatistics, Medical University of South Carolina, Charleston, SC; Geert Molenberghs, Iniversiteit Hasselt & Katholieke Universiteit Leuven; Debajyoti Sinha, Florida State University

Key Words: Bridge density, Logistic link, Longitudinal data, Non-random missingness, Shared random effects

Incomplete data is common in longitudinal studies, owing to subjects missing one or more follow up visits. When the missing-data mechanism depends on the outcome of interest, inferences that only use the complete data will no longer be valid. Shared random-effects models that uses latent variables to link the binary longitudinal response to the missing-data mechanism have been proposed, within a Gaussian framework. However, such models only describe the conditional model of the binary response given random effects and do not yield a closed-form marginal model for either the binary longitudinal response or the missing-data mechanism. We propose a shared random-effects model that, unlike the existing models, allows to preserve a logistic regression form for, at the same time, the marginal probability of the binary response, the missing-data indicator, and the conditional probability of the binary response and missing indicator given the random effects. Markov chain Monte Carlo (MCMC) technology is used for sampling from the posterior distributions of parameters and latent variables. The methodology is applied to an example of longitudinal data of binary CD4 counts.

Longitudinal Data Analysis With Both Monotone And Non-Monotone Missing Data

✦Chi-hong Tseng, UCLA department of medicine, 10940 Wilshire Blvd, Suite 1223, Los Angeles, CA 90024, *tseng.ch@gmail.com*; Robert Elashoff, UCLA

Key Words: missing data, shared parameter model, integrated likelihood

A common problem in the longitudinal data analysis is the missing data problem. Two types of missing patterns are generally considered in the statistical literatures: monotone and non-monotone missingness. Non-monotone missing data are generally caused by intermittent missed visits by study participants. Monontone missing data can be from discontinued participation, loss to follow-up and mortality. Although many novel statistical approaches have been developed to handle missing data in recent years, few methods are available to provide inferences to handle both types of missing data simultaneously. In this research, a joint shared parameter model is proposed to analyze longitudinal outcome data with both monotone and non-monotone missingness. To overcome the difficulty of high dimensional integration in the estimation of the shared parameter model, we propose to use adaptive quadrature for computational efficiency since it requires fewer quadrature points to achieve the same precision. Simulation study is carried out and a real data example from the Scleroderma lung study is used to demonstrate the effectiveness of this method.

Mixed-Effects Models for the Effects of Partially Observed Covariates on Dependent Recurrent Events Data

✦Stephen Rathbun, University of Georgia, 132B Coverdell Center, Athens, GA 30602, *rathbun@uga.edu*

Key Words: point process, ecological momentary assessment, estimating equations, smoking data

In an Ecological Momentary Assessment of smoking, volunteer smokers were asked to record each cigarette smoked on a PDA over a 16-day period prior to a designated quit date. In addition, the PDA prompted smokers to answer questions regarding mood and environment at times of randomly sampled cigarettes and at times selected according to a probability-based sampling design. A mixed effects version of a point process model is constructed to describe the relationship between the repeated cigarette smoking events and the partially-observed time-varying covariates, allowing for variation among smokers in their responses to mood and environment. Estimating equations are proposed that yield consistent estimators of model parameters without making assumptions regarding the distribution of time-varying covariates. Moreover, statistical inference remains valid even when, conditional on the random effects, event times are not realized from a Poisson point process.

Joint Modeling Of Longitudinal And Survival Data With Missing And Left-Censored Time-Varying Covariates

✦Ryan C May, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, *ryanmay@unc.edu*; Joseph G. Ibrahim, University of North Carolina; Haitao Chu, University of Minnesota, School of Public Health, Division of Biostatistics; Stephen R Cole, University of North Carolina at Chapel Hill

Key Words: Joint Modeling, Detection Limit, HIV, MACS

We propose a joint model for longitudinal and survival data with time-varying covariates subject to detection limits and ignorable intermittent missingness. The model is motivated by data from the Multicenter

Aids Cohort Study (MACS), in which HIV+ subjects have viral load and CD4 cell counts measured at repeated visits along with survival data. We model the longitudinal component of the joint model via a generalized linear mixed model (GLMM), predicting the trajectory of CD4 cell counts with viral load and other covariates. The viral load data is subject to both left-censoring due to detection limits (17\%) and ignorable intermittent missingnes (27\%). The survival component of the joint model looks at death due to AIDS, and is taken as a Cox proportional hazards model. The longitudinal and survival models are linked via the trajectory function of the GLMM, which is included in the survival component. A Bayesian analysis is conducted on the MACS data using the proposed model.

The Performance Of Unadjusted And Adjusted Medical Cost Estimators -- An Application In Colorectal Cancer

✦Yi-Ting Hwang, National Taipei University, 67, Sec. 3, Ming-Sheng E. Rd., Taipei, 104 Taiwan, *hwangyt@gm.ntpu.edu.tw*

Key Words: Inverse probability, Longitudinal data, Medical cost, PH model, Random effect model

How to sufficiently use the limited resources becomes controversial owing to aging. It is then important to know how to estimate medical cost accurately and efficiently. The naÔve estimators ignoring the unobservable data may be biased owing to dropouts. Lin (1997) suggested partitioning the study duration and then constructing the estimate by summing up the cost from each interval. Furthermore, to take into account of the unobservable data, Lin (1997) and Band and Tsiatis (2000) proposed weighted estimators that used the survival probability and uncensored probability as the weight, respectively. The medical cost may be related to many covariates. Baser (2006) suggested using the general linear model for the longitudinal data to model the partitioned cost, where a random intercept is included. This paper extends the model to a more general parametric model. Furthermore, we suggest using the survival probability as the weight adjustment which is estimated by the Cox proportional hazards model. This paper compares the performance of the unadjusted and adjusted cost estimators under various scenarios using simulations. Finally, the proposed model is implemented on a real data.

Evaluating The Effects Of Treatment Regimes In The Presence Of Drop-Outs: Application To Depression Data

✦Yenchih Hsu, University of Pittsburgh, 130 DeSoto Street, 326 Parran Hall, Pittsburgh, PA 15261, *yeh7@pitt.edu*; Adbus S. Wahed, University of Pittsburgh

Key Words: Dynamic treatment regime, Generalized estimating equations, Inverse-probability-weighting, Missing data, Two-stage designs

Studies in mental health research often employ two-stage designs to assess combinations of pharmacotherapies and psychotherapies. Outcomes in such studies often consist of repeated measurements of scores such as the 24-item Hamilton Rating Scale for Depression (HRSD) over a fixed duration of time. Since treatment is given sequentially, the eligibility of receiving one treatment assignment depends on previous treatments and intermediate outcomes. The goal is to compare different treatment regimes in two-stage longitudinal studies to find the most beneficial one for each patient. The presence of missing data is a common phenomenon in longitudinal studies (drop-outs, withdrawals, etc.). In this paper, we show how to construct proper weights to account for monotone missing and apply them in the estimating equations to draw inference for treatment regimes from two-stage longitudinal studies. Specifically, we provide consistent estimators and their asymptotic variances of the effects of treatment regimes. Large-sample properties of the proposed estimators are provided analytically, and examined through simulations. We apply our methods to a depression dataset that motivated this study.

464 Bayesian Methodology in life sciences

Section on Bayesian Statistical Science

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Bayesian Inference On The Population Size In The Capture-Recapture Model

✦Chang Xu, University of Missouri, 22 Broadway Village Dr, Apt A, Columbia, MO 65201, *cx3z9@mail.missouri.edu*; Dongchu Sun, University of Missouri

Key Words: Capture-recapture model, Reference prior

Capture-recapture models are widely used in ecology to estimate the population size. Bayesian methods are frequently used to analyze the capture-recapture models. However, one of the unsolved difficulties is to choose the appropriate prior for the population size parameter, N. In this paper we present the reference prior for N following the novel approaches proposed by Berger, Bernardo and Sun(2010). We also conduct simulations to compare the reference prior with the traditional non-informative priors, based on criteria such as posterior coverage and MSE of the Bayesian estimate.

Regression Splines For 1-1 Matched Case-Control Studies With Measurement Error

✦Nels Johnson, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060, *nels@vt.edu*; Inyoung Kim, Virginia Polytechnic Institute and State University

Key Words: measurement error, matched case-control study, Bayesian methods, regression splines

In epidemiological research, matched case-control studies are popular. Measurement error of covariates is a modeling problem that often needs to be addressed. We adapt a Bayesian method for regression splines in the presence of measurement error for use in 1-1 matched case-control studies. We also discuss some methods to address computational issues that arise when using this method.

Bayes Factors in the Presence of Population Stratification

✦Linglu Wang, Department of Statistics, George Washington University, 2140 Pennsylvania Ave NW, Washington, DC 20052, *linglu@gwmail.gwu.edu*; Qizhai Li, Academy of Mathematics and Systems Science, Chinese Academy of Sciences; Zhaohai Li, Department of Statistics, George Washington University; Gang

Zheng, Office of Biostatistics Research, National Heart, Lung and Blood Institute

Key Words: approximate Bayes factor, clustering, genome-wide association studies, odds ratio, population stratification, principal components analysis

Population stratification (PS) is a main concern of using the case-control design to detect genetic association. All methods to correct for PS have been studied in classical hypothesis testing. It is not clear, however, how the PS would affect Bayesian hypothesis testing and how to correct for it in Bayesian analysis. In this paper, we start with the investigation on how the PS would affect the asymptotic distribution of the estimate of odds ratio and its confidence interval. Using a large panel of null markers scanned in a genome-wide association studies, we apply a principle clustering analysis method to adjust for the effect of PS on the estimate of odds ratio and its asymptotic variance. Then we apply these corrected estimate and the asymptotic variance to an approximation of Bayes factor. We show in simulations that the proposed method yields an appropriate correction for the PS in Bayesian analysis and apply the results to HapMap data as an illustration.

A Bayesian Subgroup Analysis Using An Additive Model

✦Yang Xiao, University of Cincinnati, 33 Ridgewood PL, Fort Thomas, KY 41075, *xiaoyo@mail.uc.edu*; Siva Sivaganesan, University of Cincinnati; Purushottan Laud, Medical College of Wisconsin; Peter Mueller, MD Anderson Cancer Center

Key Words: additive model, Bayesian subgroup analysis, stepwise model selection, multiplicity

Subgroup analysis in clinical trials examines whether a treatment efficacy differs among the sub-population defined by one or a set of baseline characteristics, which will potentially influence future health decision making and research for clinical investigators. However, difficult issues such as multiplicity adjustment, sample size and power need to be addressed and reported correctly. We extend a Bayesian approach in Sivaganesan et. al, by using a single additive model with covariates each at two levels to allow for the treatment-covariate interaction. Posterior probabilities along with carefully chosen threshold values are used to decide on subgroup effects.We account for multiplicity in the Bayesian sense by the use of suitable prior distributions for the inclusion parameter(s), and in the frequentist sense by control of Type I error rate. Frequentist operating characteristics of the approach is investigated under various scenarios considering various error rates specific to subgroup analysis.

An Extension Of Continual Reassessment Method

✦Xiaobin Yang, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249 USA, *xiaobin.yang@utsa.edu*; Keying Ye, University of Texas at San Antonio

Key Words: Polychotomous toxicity response, Continual reassessment method, Latent variable, Gibbs sampler

Continual reassessment method (CRM) is widely used in clinical trials for fnding the maximum tolerated dose (MTD), especially in the case of dichotomized (or binary) toxicity response. By introducing a continuous latent variable we extend the CRM in the case of polychotomous (or multinomial) toxicity response. The polychotomous toxicity grades are determined according to the disjoint intervals in which the latent variable falls. We first introduce the latent variables in the case of dichotomized toxicity response. Compared with the traditional CRM, the method with latent variable is applicable and comparable based on the simulation studies. Then we apply the method with latent variable in the case of polychotomous toxicity response. The joint posterior distribution and the fully conditional distributions of the parameters are obtained. Furthermore, simulation studies using the Gibbs sampler in the case of polychotomous toxicity response will be given.

The K-Zig: A Flexible Model For Zero-Inflated Counts

✦Souparno Ghosh, Duke University, , *sg147@stat.duke.edu*; Alan E Gelfand, Department of Statistical Science; James S. Clark, Duke University; Kai Zhu, Duke University

Key Words: Zero-inflated models, logit link, Posterior predictive loss function, FIA data

In many applications involving count data, one comes across data generating processes yielding significantly high number of zeros. Zero-inflated Poisson(ZIP) and zero-inflated negative binomial (ZINB) models are generally used to deal with such situations. However, these traditional models require very large amount of data to estimate the parameters accurately when one encounters extremely high proportion of zeros, say more than 80\%. In other words, when sample size is moderate and/or covariate information is weak, the ZIP or ZINB models are not flexible enough to handle such high proportion of zeros. To redress this problem we propose the k-ZIG model that allows more flexible modeling of zero-inflation and non-zero counts. The model is fitted within a Bayesian framework. The methodology is illustrated with simulated data examples as well as forest seedling data obtained from the Forest Inventory and Analysis National Program.

465 Bayesian Modeling and Computation ■●

Biometrics Section, ENAR, Section on Bayesian Statistical Science, Section on Statistics in Epidemiology

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Exact Tests Of Hardy-Weinberg Equilibrium A La Fisher In Multi-Allelic Biomarkers And Algebraic Statistics

✦Marepalli B Rao, University of Cincinnati, 106 Kettering Lab, Dept. Environmental Health, Cincinnati, OH 45267, *marepalli.rao@uc.edu*; Subramanyam Kasala, University of North Carolina at Wilmington; Marepalli B Rao, University of Cincinnati

Key Words: Exact Test, Hardy-Weinberg Law, MCMC, p value, Permutation test, Alleles

In the context of large contingency tables with sparse data and/or highly imbalanced entries, algebraic statistics methodology is successfully implemented to conduct a test of independence a la Fisher. In this presentation, we want to rope in algebraic statistics to develop an exact test of Hardy-Weinberg Equilibrium a la Fisher in multi-allelic biomarkers.

The test is a combination of fibers, Markov bases, and MCMC. The key idea is to look at the fiber of all possible genotype data sets with allele frequencies the same as those of the given genotype data set. It turns out that the fiber can be characterized as the collection of all symmetric matrices with non-negative integer entries, even diagonals, and row sums the same as the allele frequencies of the given genotype data set. Even for moderate number of alleles and sample sizes, the fiber could be humongous. We will construct a Markov base which with the given data set can generate the entire fiber. The size of a Markov base is much, much smaller than the size of the fiber. We then develop an MCMC algorithm to estimate the p value under the equilibrium.

Fitting Social Network Models Using Varying Truncation Stochastic Approximation Mcmc Algorithm

✦Ick Hoon Jin, Texas A&M University, 301 Holleman Dr. E Apt. 822, College Station, TX 77840, *kentjin@stat.tamu.edu*

Key Words: Exponential Random Graph Model, Stochastic Approximation Markov Chain Monte Carlo, Model Degeneracy, Trajectory Averaging, Monte Carlo MLE, Social Network

The exponential random graph model (ERGM) plays a dominant role in social network analysis. However, the current methods, such as Monte Carlo MLE and stochastic approximation, often suffer from the model degeneracy problem in fitting ERGMs, rendering failures in parameter estimation. In this paper, we introduce a varying truncation stochastic approximation Markov chain Monte Carlo (SAMCMC) method for estimating the parameters of ERGMs. The varying truncation mechanism enables SAMCMC to overcome the model degeneracy problem. Under mild conditions, we show that the resulting estimator is consistent, asymptotically normal, and asymptotically efficient. The SAMCMC method is illustrated using a variety of social networks. The numerical results indicate that SAMCMC can significantly outperform the Monte Carlo MLE and stochastic approximation methods. For the ERGMs which consist of basic Markovian statistics, the Monte Carlo MLE and stochastic approximation methods often fail due to the model degeneracy, while SAMCMC still works well. For the ERGMs which do not suffer from the model degeneracy, SAMCMC can work equally well as or better than the pre-existing methods.

Massively Univariate Regression Accounting For Random Regressors

✦Xue Yang, Electrical Engineering, Vanderbilt University, 2301 Vanderbilt Pl., PO Box 351679 Station B, Nashville, TN 37235, *xue.yang@vanderbilt.edu*; Carolyn Lauzon, Vandebrilt University; Ciprian Crainiceanu, Johns Hopkins University; Brian Caffo, Johns Hopkins Department of Biostatistics; Susan M. Resnick, National Institute on Aging, National Institutes of Health; Bennett Landman, Vanderbilt University

Key Words: massively univariate regression, model II regression, GLM

Massively univariate regression with the general linear model (GLM) is an essential method for quantitative interpretation of multi-dimensional medical images, such as magnetic resonance imaging (MRI) or positron emission tomography (PET). These mappings enable inferences on image appearance (e.g., tissue characteristics) conditioned on a set of non-image based regressors (e.g., treatment regimen). With modern technology, single subject multi-modal imaging data are increasingly common and regression between two images (e.g., MRI vs. PET) has become highly desirable. Although software tools have been extended to support images as regressors they do not account for randomness in the regressor image or model the joint distribution. Model II regression could address this problem but a general maximum likelihood (MLE) solution is ill conditioned. Here, we combine Model II regression with data-derived prior variances to regularize the MLE solution, producing a model that is inverse consistent. This approach shows substantial improvements in fit and increased statistical power compared to OLS, and provides a logical framework for exploring relationships in multi-modal image analysis.

Locally Optimal Sampler

✦Ting-Li Chen, Academia Sinica, Institute of Statistical Science, 128 Academia Rd, Sec. 2,, Taipei, 115 Taiwan, *tlchen@stat.sinica.edu.tw*; Shang-Ying Shiu, National Taipei University

Key Words: Markov chain Monte Carlo, Gibbs sampler, Metropolis-Hasting sampler, asymptotic variance

Let A be a finite space and p be an underlying probability on A. For any real-valued function f defined on A, we are interested in calculating the expectation of f under p. Let X1, X2, X3, ... be a Markov chain generated by some transition matrix P with invariant distribution p. The time average, the summation of f(Xk) is a reasonable approximation to the expectation. Which matrix P minimizes the asymptotic variance of the time average? The answer depends on f. First, we identify the set of P's that minimize the average asymptotic variance, averaged with respect to a uniform distribution on f. Then we propose a new Markov chain Monte Carlo algorithm by updating locally using one of the optimal transition matrices. Our simulation results show that the new algorithm Locally Optimal Sampler has substantially lower variances than either the Gibbs sampler or the Metropolis-Hastings sampler.

Measurement Error and Misclassification in Covariates: Impacts on the Main and Interaction Effects and the Bayesian Adjustment

✦Shahadut Hossain, United Arab Emirates University, Department of Statistics, UAE University, P.O. Box 17555, Al Ain, United Arab Emirates, *shossain@uaeu.ac.ae*

Key Words: Measurement Error, Misclassification, Errors-in-variables, Exposure Model

Measurement error in quantitative factors or misclassifications in categorical factors or both are unavoidable in most epidemiological studies. If ignored, such measurement errors and/or misclassifications can seriously distort the estimated effects of the risk factors on the disease outcome of interest. Thus, adjustments for measurement errors and/or misclassifications are necessary to have accurate estimates of the effects of the suspected risk factors, and their interactions on the disease risks. In this talk, I will discuss a general Bayesian flexible parametric framework for eliminating the deleterious impacts of measurement errors and/or misclassification while estimating the main and the interaction effects in logistic regression model. The application of the proposed flexible method will be illustrated with reference to gene-environment interaction studies. Being a parametric framework, the method is unified and thus can be applied to any member of the generalized linear models, and to the non-linear regression models as well.

A Bayesian Model Averaging Approach For Estimating The Relative Risk Of Mortality Associated With Heat Waves In 105 U.S. Cities

✦Jennifer F. Bobb, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St., Baltimore, MD 21205, *jenniferfederbobb@gmail.com*; Francesca Dominici, Harvard School of Public Health; Roger D. Peng, Johns Hopkins School of Public Health

Key Words: Climate change, Generalized additive models, Model uncertainty, Time series data

Estimating the risks heat waves pose to human health is a critical part of assessing the future impact of climate change. We propose a flexible class of time series models to estimate the relative risk of mortality associated with heat waves and conduct Bayesian model averaging (BMA) to account for the multiplicity of potential models. Applying these methods to data from 105 U.S. cities for the period 1987-2005, we identify cities having a high posterior probability of increased mortality risk during heat waves, examine the heterogeneity of the posterior distributions of mortality risk across cities, assess sensitivity of the results to the selection of prior distributions, and compare our BMA results to a model selection approach. We find that no single model best predicts risk across cities, and that for some cities heat wave risk estimation is sensitive to model choice. While model averaging leads to posterior distributions with increased variance as compared to statistical inference conditional on a model obtained through model selection, we find that the posterior mean of heat wave mortality risk is robust to accounting for model uncertainty over a broad class of models.

466 Statistical Methods for High-Throughput and High-Dimensional Data

Biometrics Section, ENAR, Section on Statistics in Epidemiology, Section on Statistical Computing

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Preliminary Test Estimation Procedures In High Throughput Screening Assays

✦Changwon Lim, National Institute of Environmental Health Sciences, 111 T. W. Alexander Dr, RTP, NC 27709, *limc2@niehs.nih.gov*; Pranab K Sen, University of North Carolina at Chapel Hill; Shyamal D Peddada, National Institute of Environmental Health Sciences

Key Words: Dose-response study, False discovery rate (FDR), Heteroscedasticity, Hill model, M-estimation procedure, Nonlinear regression model

Quantitative high throughput screening (qHTS) assays use cells or tissues to screen thousands of compounds in a short period of time. Data generated from qHTS assays are then fitted using a nonlinear regression model and decisions regarding toxicity of a chemical are made using the estimates of the parameters of the model. For such data sets, the error may be homoscedastic or heteroscedastic. Because thousands of compounds are evaluated in qHTS assays based on the estimates of the parameters of nonlinear models, it is important to apply an estimation procedure which is robust to the underlying variance structure as well as robust in the presence of outliers in the data. In this talk we describe a preliminary test estimation (PTE) based methodology for drawing inferences regarding the parameters of a Hill model with application to qHTS assays. Performance of the PTE based methodology, in terms of false discovery rate (FDR) and power, is evaluated using a simulation study mimicking a real qHTS data. Results indicate that the PTE based methodology achieves substantial reduction in FDR while not losing much power. The proposed methodology is illustrated using a data set from NTP.

An Information Ratio Criterium For Integration Of High Throughput Data Sources With Applications To Quality Control And Novelty-Detection In Pluripotent Stem-Cell Cultures

✦Bernhard Michael Schuldt, AICES Graduate school, RWTH Aachen, Aachen, 52062 Germany, *schuldt@aices.rwth-aachen.de*; Michael Lenz, AICES Graduate school, RWTH Aachen; Johanna Goldmann, Institut f¸r Biochemie, Freie Universit‰t Berlin; Franz-Josef M¸ller , Zentrum f¸r Integrative Psychiatrie ; Jeanne Loring, Center for Regenerative Medicine, The Scripps Research Institute; Andreas Schuppert, AICES Graduate school, RWTH Aachen

Key Words: microarray, DNA-Methylation, next-generation sequencing, novelty detection, quality control

Human pluripotent stem cells have become a focus of academic and industry researchers alike. Currently, the quality of such in vitro preparations is assessed mainly with qualitative, low-content assays such as the teratoma assay. Alternatively high-thoughput methods can be used to generate data-driven models to evaluate pluripotency, to assess quality and to detect novel features with biological implications in the data . Technologies such as mRNA-, SNP- and methylation arrays as well as next-generation deep sequencing have been proposed for this task. To establish an integrative quality control framework in a fast developing field it is key to understand how information is preserved with new profiling technologies. We use the information ratio (IR), a measure that compares gene-wise measurements to large scale patterns, to develop a classification of various quality control and novelty detection tasks. We evaluated the IR heuristic on a large, heterogeneous stem cell data set and developed practical guidelines for an integrated quality control workflow that can be easily extrapolated to data generated in other laboratories or can be applied to other fields.

Noise Reduction in Genome-Wide Perturbation Screens Using Linear Mixed-Effect Models

✦Danni Yu, Purdue University, 3461 Brixford Ln, West Lafayette, IN 47906 USA, *dyu@stat.purdue.edu*; John Danku, Purdue University; Ivan Baxter, USDA-ARS Plant Genetics Research Unit, Donald Danforth Plant Science Center; Sungjin Kim, Cornell University ; Olena K. Vatamaniuk, Cornell University ; David E. Salt, Purdue University; Olga Vitek, Purdue University

Key Words: statistical analysis, linear mixed-effect models, noise reduction, genome-wide perturbations, high-througput screening, data normalization

High-throughput perturbation screens measure the phenotypes of thousands of biological samples under various conditions, which are subject to substantial biological and technical variation. Meanwhile, it is often impossible to include a large number of replicates and to randomize the order of the replicates in high throughput screens. Distinguishing true changes in the phenotype from stochastic variation in such experimental designs is extremely challenging, and requires adequate statistical methodology. We propose a statistical modeling framework that is based on experimental designs with at least two controls profiled in the experiments, and a normalization and variance estimation procedure with linear mixed-effects models. We evaluate the framework using three comprehensive screens of yeast: 4940 gene-deletion haploid mutants, 1127 gene-deletion diploid mutants, and 5798 gene-overexpression haploid mutants. The proposed approach can be used in conjunction with practical experimental designs, allows extensions to alternative experimental workflows, enables a sensitive discovery of biologically meaningful changes, and strongly outperforms the existing noise reduction procedures.

Mixture Similarity Peak Alignment For Two-Dimensional Gas Chromatography Mass Spectrometry Based Metabolomics

✦Seongho Kim, University of Louisville, 485 E. Gray ST., Louisville, KY 40222, *s0kim023@louisville.edu*; Aiqin Fang, University of Louisville; Bing Wang, University of Louisville; Jaesik Jeong, Indiana University School of Medicine; Xiang Zhang, University of Louisville

Key Words: Gas Chromatography, Metabolomics, Peak Alignment, Retention time, Mass Spectrometry

Comprehensive two-dimensional gas chromatography mass spectrometry provides more accurate information about metabolite retention times and mass spectrum. However, there is always a shift of retention times in the two GC columns due to the differences in temperature, pressure, matrix effects, and stationary phase degradation, making it difficult to compare metabolic profiles. The existing peak alignment algorithms use the peak distance and the spectra similarity sequentially and require predefined either distance-based window or spectral similarity-based window. To overcome these limitations, we developed a mixture similarity peak alignment algorithm by employing the peak distance and the spectral similarity measures simultaneously without any variation windows. We further examined the effect of the four different distance measures, Euclidean, Maximum, Manhattan, and Canberra distances. The performance of our proposed algorithm was compared with the existing methods on the two real data. Our analysis showed that Canberra distance performed better than other distances and the proposed mixture similarity peak alignment algorithm prevailed against all literature reported methods.

Spatial Clustering For Biomarker Discovery In Nuclear Magnetic Resonance-Based Metabolomics Data

✦Xu Han, Baylor College of Medicine, Neurological Research Institute,, 1250 Morusund Street, Houston, TX 77030, *Xhan@bcm.tmc.edu*; Genevera I. Allen, Baylor College of Medicine & Rice University; Mirjana Maletic-Savatic, Baylor College of Medicine

Key Words: spatial clustering, biomarker discovery, pattern recognition, metabolomics

NMR spectroscopy technology produces high-dimensional metabolomic profiles that are mixtures of several latent metabolites often resonating at multiple chemical shifts. To aid in biomarker discovery, the spectrum must be grouped into candidate metabolite regions. We propose to group these variables into spatial clusters, or contiguous sets of variables exhibiting high correlation. The average pair-wise correlation of these contiguous sets is compared to a permutation null distribution obtained by permuting the order of the variables. The optimal threshold used to determine the extent of the spatial clusters is obtained by comparing the average within-cluster correlation to that observed in the original spectrum. We present simulated and real-data examples illustrating the utility of our approach in discovering candidate metabolites from NMR spectroscopy data.

Bayesian Functional Classification And Visualization Of Neurons For Brain Decoding

✦Duchwan Ryu, Georgia Health Sciences University, 120 15th Street, AE - 1005, Augusta, GA 30912, *dryu@georgiahealth.edu*

Key Words: Bayesian nonparametric functional analysis, Bayesian partition, dimension reduction

Neurons react and memorize on external stimuli with spikes on raster data. We first characterize each neuron through a Bayesian nonparametric functional analysis based on its frequencies of spikes, and then apply a Bayesian classification method hierarchically to establish neural cliques which are network-level coding units in brain. Because the hierarchical neural cliques depend on the imposed stimuli, we visualize them on dimension reduced stimuli space. With hippocampal data of mice, we demonstrate activities of large numbers of neurons in neural circuits during memory acquisition, consolidation and retrieval for mnemonic startling episodes.

Shrinkage, Glm-Ridge And Penalty Estimators In Generalized Linear Models

✦Md Shakhawat Hossain, Alabama A & M University, Dept of Mathematics; P.O. Box 326, 4900 Meridian Street, Huntsville, AL 35762, *shakhawat.hossain@aamu.edu*; Ejaz Syed Ahmed, University of Windsor; Kjell Doksum, University of Wisconsin

Key Words: Shrinkage, Ridge, LASSO

We consider the estimation problem for the generalized linear models which may have a large collection of potential predictor variables and some of them may not have influence on the response of interest. In this situation, selecting the statistical model is always a vital component in estimation procedure. In the context of two competing models, we demonstrate the relative performances of shrinkage and glm-ridge estimators based on the asymptotic analysis of quadratic risk functions and it is found that the shrinkage and glm-ridge estimators outperform the maximum likelihood estimator (MLE) uniformly. For comparison purpose, we also consider the L1-penalty type estimator for generalized linear models. A Monte Carlo simulation study has been conducted in comparison of shrinkage, glm-ridge and penalty type estimators with respect to MLE through relative efficiency. This comparison shows that shrinkage and glm-ridge methods perform better than the L1-penalty type estimation method when the dimension of the restricted parameter space is large. A real data set analysis is also presented to compare the suggested methods.

467 Statistical Methods ■●

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

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Evaluating Energy Efficiency Of Demand-Side Management Programs

✦Cha-Chi Fan, Department of Energy, , *ChaChi.Fan@eia.gov*

Key Words: energy efficiency, demand response, demand-side management, nonequivalent groups

Energy efficiency is defined differently for various purposes, making calculating end-user energy efficiency controversial in the context of demand-side management. One fundamental issue is that pure measurements of energy efficiency cannot be obtained without influence from confounding factors (i.e.: weather or equipment maintenance). Thus, several efficiency index decomposition analysis methods have been used to remedy the lack of uncontaminated data. However, most index decomposition analysis is done at an industrial level, which provides a better-controlled environment than that of the end user's level. Required information at an end user's level is much more difficult to collect making index decomposition analysis impractical. As the result, evaluation of energy efficiency associated with demand-side management programs becomes subjective. An assessment of whether the energy savings and actual peak load reduction varies by instituting demand-side management programs or not, and how different types of customers respond to them, is proposed in this study by using supplemental statistical analysis to reduce the influence of unwanted confounders.

Methods Of Estimation In Random Effects Meta-Regression

✦Abera Wouhib, NCHS/CDC, 3311 Toledo Road Rm # 3113, Hyattsville, MD 20782, *awouhib@cdc.gov*

Key Words: Meta-regression model, Random effects, Heterogeneity parameter

The term meta-regression often refers to a linear models of effect size with covariates. Unlike common versions of regression analyses, which include available record-level data on outcomes and covariates, meta-regressions may consist study-level covariates in addition to record-level covariates. Random effects meta-regression models account for non-systematic differences among study outcomes which cannot be explained by sampling variability alone. We demonstrate a new estimation method using a random effect meta-regression model along with other methods of meta-analysis to estimate effect size of interest and the heterogeneity component of the variance using a simulation study. In the simulation study, we draw samples from a stochastic model that closely emulates the design of health surveys. We select the estimators of effect size which are best in terms of minimizing mean-square-error and bias. We then apply these estimators to data from health surveys and explore the causes of heterogeneity in effect sizes by including covariates at the study-level, record-level or both. Finally, we compare statistical methods, and discuss the limitations and drawbacks of these methods.

Maximum Likelihood Skew-Normal Factor Analysis

✦Beverly Jane Gaucher, Texas A&M University, 800 Marion Pugh, Apt. 2213, College Station, TX 77840 USA, *beverlygaucher@msn.com*

Key Words: skew-normal, factor analysis, maximum likelihood

This research explores factor analysis applied to skewed distributions for a general skew model, the skew-elliptical model and the skew-normal model for all sample sizes. The skewed models are formed using selection distribution theory, which is based on Rao's weighted distribution theory. The models assume the observed variable of the factor model is from a skew-normal distribution by defining the distribution of the unobserved common factors skew-normal and the unobserved unique factors Gaussian noise. The skew-normal factor analysis model's log-likelihood is derived from which the maximum likelihood factor loading estimates are calculated.

Optimal Combination Of Structured Classifications

✦Mark van der Loo, Statistics Netherlands, Henri Faasdreef 312, The Hague, 2492 JP The Netherlands, *mark.vanderloo@gmail.com*

Key Words: stratification, hierarchical classification, optimization, simulated annealing, heuristic methods

Dividing populations in homogeneous subsets is an important method to improve the quality of statistical statements. In official statistics, one is often confronted with demographic or economic (hierarchical) classifications which define meaningful but not necessarily homogeneous subsets. Multiple classifications can yield unacceptably small subpopulations so in practice, subject matter specialists have to combine classes to obtain large and homogeneous enough subpopulations, taking into account the (hierarchical) structure of multiple classifications. Here, I present two automated methods for optimal class-combination which are able to take account of multiple structured classifications. The first method is based on a tree search algorithm which heuristically searches for combinations that meet a set of user-defined requirements. The second method is based on simulated annealing and optimizes a global cost function. An application to Dutch turnover data is also demonstrated. The methods have been implemented in R and will be made available via the CRAN archive as the stratifier package.

Statistical Methods In Compensation Discrimination Analysis

✦Bogong Timothy Li, OFCCP, 200 Constitution Ave., NW, Washington, DC 20210 USA, *bli@dol.gov*

Key Words: Compensation analysis, Multiple regression, Categorical analysis, EEO

Statistical methods are critical in identifying pay patterns that may be a result of discriminatory compensation practices. Supreme Court has decided a number of cases specifically relied on such statistical evidence. Federal laws also require statistical analysis in order to justify certain findings. Nevertheless many disagreements exist among legal and statistical professionals in how to interpret statistical results, and the difference between methods, such as those from multiple regressions, categorical analysis or small sample methods etc. In addition to a general summary of most often used methods by federal and state EEO agencies, the paper intends to provide a coherent interpretation of statistical methods, reconcile the differences between various methods and provide some specific approaches to certain difficult areas of analysis.

The Role Of Legal Orientation Program In Affecting Immigration Court Case Processing Times

✦Zhifen Cheng, Vera Institute of Justice, 233 Broadway, Fl 12,
New York, NY 10279, *zcheng@vera.org*

Key Words: immigration, court, history event

Since 2005, the Vera Institute of Justice has been contracted by the Department of Justice's Executive Office for Immigration Review (EOIR), the nation's immigration courts, to manage and evaluate its Legal Orientation Program which is a multi-site program providing legal rights information to non-citizens detained by the DHS. The evaluation of the LOP, aiming to assess its impact on case outcomes and the immigration courts, found that cases for LOP participants took fewer days to go through the immigration courts than those for detainees not participating in the program. As a follow-up to that finding, we conducted a history event analysis of the immigration court data and the LOP program data, taking into consideration the fact that analysis of completed cases alone would cause estimation bias. The results confirmed statistically that LOP accounted for shorter case processing times for detained cases. Because of the immigration court system's relatively brief history, there was no rigorous prior research focused on case processing times of immigration court cases. Our study contributed to the development of the field by analysis beyond basic descriptive statistics.

468 Health and Medical Risk

Section on Risk Analysis, ENAR, Section on Health Policy Statistics, Section on Statistics in Defense and National Security, Section on Statistics in Epidemiology, Section on Quality and Productivity, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Theory Of Individual Health Histories And Dependent Competing Risks

Igor Akushevich, Duke University; ✦Konstantin Arbeev, Duke University, Durham, NC 27708 USA, *ka29@duke.edu*; Svetlana Ukraintseva, Duke University; Anatoliy Yashin, Duke University

Key Words: cause-specific mortality risks, longitudinal data, risk of disease onset, observational plans, age trajectories, stochastic differential equations

Data on individual health histories, age trajectories of physiological variables and cause-specific mortality allow for studying joint evolution of health and physiological states and their effects on mortality risk. Generally cause-specific mortality risks are mutually dependent. Here we present a competing risk model with a weaker assumption of conditional independence of cause-specific time to death, given a stochastic process with two mutually dependent continuous and jumping components. The jumping component describes fast changes in health status and the continuous one describes slower individual physiological aging. Formulation of the exact model based on the Kolmogorov equations for conditional density, the Gaussian approximation, and the parameter estimation in different observational plans in longitudinal measurements are discussed and properties of the model estimation are investigated in simulation study. Application to the Framingham Heart Study resulted in evaluation of a system of indices representing aging related changes and indicators of health. The results suggest the importance of avoiding the marginal independence assumption when appropriate data are available.

Pregnancy Delivery Cost Analysis And Risk Adjustment For Hospital

✦Jie Gao, CareFirst BCBS, MD 21231, *jaojie@gmail.com*

Key Words: risk factors, pregnant cost, general linear model, risk adjustment, reduce cost, healthcare cost

Although we know that "the estimated cost of delivery alone is $6,000 - $8,000 for a normal pregnancy and the cost increases if it is a high risk pregnancy"1(American Pregnancy Association), there is little information in the literature regarding the risk factors that increase the average medical expenditures generated in the hospital delivery. Competition among the delivery hospitals would always help to improve the quality of care for patients and potentially lower the cost from Insurance Companies. So the efficiency and fairness of the evaluation are important. The general linear regression models are used to capture potential risk factors contributing to the facility delivery cost are constructed, which enables us to apply a risk adjustment technique to fairly evaluate facilities with different illness burdens of patients and thus identify the potential to reduce the costs and improve the quality.

Abo/Rh Blood Types Affect The Adequacy And Resilience Of The U.S. Blood Supply

✦Arianna Simonetti, U.S. Food and Drug Administration, CBER, 1401 Rockville Pike, HFM 210, Rockville, MD 20852, *Arianna.Simonetti@fda.hhs.gov*; Mark O Walderhaug, U.S. Food and Drug Administration

Key Words: Blood Supply, Stock and Flow Model, ABO blood groups, Rh type

Understanding how much blood is available daily is of extreme importance to enable effective planning of strategies to minimize collection and donation disruptions to the blood supply in case of national emergencies. We developed a Stock-and-Flow simulation model of the U.S. Blood Supply to obtain quantitative estimates with uncertainty of blood availability disaggregated by ABO/Rh blood types. Modeling the blood supply by blood groups is essential in order to capture specific dynamics of the system as, for example, shortage of some rare blood types may have important public health impacts. We generated ABO/Rh type group predictions based on donation patterns and U.S. phenotype prevalences and implemented rules for blood use that allow for compatible blood to be selected and removed from inventory when the exact match blood type is not available during the simulation of blood demand. We observed variation in reduction up to 60% in the overall supply and investigated also for each blood group in case of a decrease in the total amount of blood available to the system and in the supply of specific blood types. We explored what conditions lead to a major loss of certain types of blood.

The Effect Of Hazard Assumptions On Split Selection Criteria And Predictive Error In Survival Trees

✦Brian White, University of South Florida, , *bdwhite@mail.usf.edu*

Key Words: Bootstrap, Brier Score, Data Mining, Prediction Error, Survival Analysis, Survival Trees

Many survival analysis techniques assume constant and proportional hazards although for many real data sets, these assumptions do not hold. Survival trees is one such technique commonly used to assign risks to individuals based on their personal covariates. We evaluate split selection criteria for survival trees when these assumptions are violated and propose new criteria that do not rely on such assumptions. We generate multiple simulated data sets under various scenarios to evaluate the predictive accuracy for the proposed techniques and compare them to commonly used techniques. In medical applications, a typical goal might be to define risk groups of patients at the time of treatment thus we employ our methods to a real data set from a clinical study of patients treated for breast cancer. We find that the prediction error of survival trees depends on the choice of split selection criterion as well as the measure used to estimate it. Our results are affected by the degree of censoring, nonconstancy, and nonproportionality of the hazards, thus the best choice may depend on characteristics of the data set in addition to any available a priori assumptions regarding the hazards.

Fitting Lin & Ying'S Additive Hazards Regression Model Using A Weighted Least Squares Estimator For Poisson Regression

✦M Brent McHenry, Brisol Meyers Squibb, , *brent.mchenry@bms.com*; STUART Lipsitz, Division of General Medicine, Brigham and Women's Hospital, Boston, MA; Debajyoti Sinha, Florida State University

Key Words: Additive Hazards, Survival Analysis, Weighted Least Squares, Computational, Poisson Linear Model, Method-of-moments

For failure time outcomes, modeling the hazard rate as an exponential function of covariates is by far the most popular. However, additive hazard rate regression models, in which the hazard rate is modeled as a linear function of the covariates, have received attention in the last few decades. If the distribution of the failure time is exponential, Aitkin et al. (2005) showed that the maximum likelihood estimates (MLE) of the additive hazard rate regression model can be obtained using a Poisson linear model. Instead of maximum likelihood, here we propose a weighted least squares (WLS) method-of-moments estimator to consistently estimate the additive hazard regression parameters. The WLS estimates are obtained by a slight modification of the Poisson additive regression estimating equations. The approach can also be used for a piecewise exponential, and we show that, as the intervals widths for the piecewise exponential converge to 0, e.g., each interval contain only one failure, the WLS estimates will be identical to the Martingale based method of moments estimate proposed by Lin and Ying (1994).

Some Measures Of Comparing Risk In Clustered Epidemiological Data

✦Tasneem Zaihra, UNB-SJ, 315 HAZEN HALL, UNB-SJ, SAINT JOHN, NB E2L4L5 CANADA, *allaboutstatistics@gmail.com*

Key Words: Cluster, Risk Difference, Risk Ratio, Relative Risk Difference, Overdispersion, Epidemiological Indices

In this study, we compare risk between two groups, where each group consists of some clusters and in each group there is a risk. The response of members of each cluster is binary and also members of the same cluster tend to have similar (correlated) responses. We discuss interval estimation of three risk measures used to compare the risk (proportion of response) of binary outcomes between the two groups namely, the risk difference (RD), the risk ratio (RR), the relative risk (RED). We further discuss situations where one risk measure is preferred over other. For instance, RR quantifies the strength of the association between a given disease and suspected risk factor and is widely used in toxicological studies. RD is preferred when we want to determine the relative order of importance of disease in terms public health issues as the level of rates can be lost in using RR and RED is preferred to see effectiveness of some intervention dealing with public health issues where one has to allocate limited resources while making public health policies.

469 Topics in Model and Variable Selection ■●

Section on Statistical Computing, Section on Statistical Graphics

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

A New Approach to Estimate the Critical Constant of Selection Procedures

✦MIN LI, CALIFORNIA STATE UNIVERSITY, SACRAMENTO, 6000 J STREET, SACRAMENTO, CA 95819-6088, *limin@csus.edu*; JACK E CHEN, BASF Corporation

Key Words: ranking and selection, sample size allocation, simulation

A solution to the ranking and selection problem of determining a subset of size m containing at least c of the v best from k normal distribution has been developed. The best distributions are those having, for example, (i) the smallest means, or (ii) the smallest variances. This paper reviews various applicable algorithms and supplies the operating constants needed to apply these solutions. The constants are computed using a histogram approximation algorithm and Monte Carlo integration.

Model Selection Via Standard Error Adjusted Adaptive Lasso

✦Wei Qian, University of Minnesota, Twin Cities, 313 Ford Hall, 224 Church Street S.E., Minneapolis, MN 55108, *weiqian@stat.umn.edu*; Yuhong Yang, University of Minnesota, Twin Cities

Key Words: BIC, model selection consistency, solution path, variable selection

The adaptive lasso is a model selection method shown to be both consistent in variable selection and asymptotically normal in coefficient estimation. The actual variable selection performance of the adaptive lasso depends on the weight used. It turns out that the weight assignment using the OLS estimate (OLS-adaptive lasso) can result in very poor performance when collinearity of the model matrix is a concern. To achieve better variable selection results, we take into account the standard errors of the OLS estimate for weight calculation, and propose two different versions of the adaptive lasso denoted by SEA-lasso and NSEA-lasso. We show through numerical studies that when the predictors are highly correlated, SEA-lasso and NSEA-lasso can outperform OLS-adaptive lasso under a variety of linear regression settings while maintaining the same theoretical properties of the adaptive lasso.

Model Selection For Multinomial Responses Via Association Rules Analysis

✦Pannapa Changpetch, Pennsylvania State University, 326 Thomas Building, Pennsylvania State University, State College, PA 16802, *puc121@psu.edu*; Dennis Lin, Pennsylvania State University

Key Words: Multinomial logit model, Association rules analysis, Model Selection

In this paper, we propose a model selection procedure for multinomial responses by implementing association rules analysis. In classical way, multinomial logit model building usually consider only main effects and ignore interactions. Here, we develop a model selection framework that is particularly important for applications in which interactions are present. We do this by (1) find (low-order and high-order) interactions among input variables (via association rules analysis) (2) select potential interactions to be considered in model building (3) convert these potential interactions into new dummy variables, and (4) perform variable selection among all input variables and the newly created dummy variables (interactions) to build up the optimal multinomial logit model.

Boosting Threshold Autoregressive Models

✦Qing Mai, University of Minnesota, 313 Ford Hall, 224 Church ST SE, Minneapolis, MN 55455, *maixx034@umn.edu*; Hui Zou, University of Minnesota

Key Words: Threshold autoregressive model, Boosting

Due to its interpretability, threshold autoregressive models(TAR) are widely used in economics, epidemiology, sociology, psychology and other fields. However, the current method of estimation for TAR involves assuming a certain form of the threshold and searching for an estimate exhaustively. Hence, the estimation is vulnerable to model misspecification and is expensive in computation. We propose to employ boosting to estimate TAR models. The proposed method does not require assumptions on the form of the threshold and is computationally efficient. Numerical studies will be presented to illustrate the performance of the new method.

Brain Magnetic Resonance Image Segmentation Via Mixture Model Averaging

✦Yichen Qin, Johns Hopkins University, 100 Whitehead hall, 3400 North Charles Street, Baltimore, MD 21218, *yqin2@jhu.edu*; Kathryn I. Alpert, Northwestern University; Carey Priebe, Johns Hopkins University; Lei Wang, Northwestern University

Key Words: Gaussian mixture model, Magnetic resonance images segmentation, Bayesian information criterion, Model averaging, Model selection

In this article, we show an application of the mixture model for the segmentation of brain Magnetic Resonance Images (MRI). The goal is to segment the brain MR images into three parts --- cerebrospinal fluid, gray matter and white matter, which will bring us more evidence on the association between the abnormality of the brain structure and Alzheimer disease. This application is built on the methodology discussed in Priebe et al. 2006 and Lee et al. 2008, but with modifications on the design of the classifier. Instead of estimating the mixture model complexity, we propose an ensemble learning approach --- model averaging --- so as to get more robust estimation. We first fit multiple mixture models to the data with model complexity specified ex ante, and then use Bayesian Information Criterion (BIC) to weight average all the models to get more robust segmentation results. The proposed methodology is well explained and is also validated in the application of segmenting real brain cingulate gyrus MRI data. We also show the comparison between the model averaging results and single model results in terms of robustness.

Identification Of Significant Features In High-Dimensional Data Using Supervised Principal Components

✦Flora Mulkey, University of North Carolina, Chapel Hill, 10 Red Bluff Ct., Durham, NC 27713, *stats151@gmail.com*; Eric Bair, University of North Carolina-Chapel Hill

Key Words: Feature Selection, Microarray, High-Dimensional Data, Semi-Supervised Methods, Genetics

We consider the problem of identifying features that are associated with an outcome of interest in high-dimensional data. This problem is particularly relevant in the analysis of DNA microarray data and genome-wide association data, where measurements from thousands of genes can be evaluated to further elucidate the genetic causes of complex diseases. We show that a method called supervised principal components can successfully identify features associated with a specific outcome more accurately than conventional methods. It is particularly effective in situations where the outcome variable is measured with a high degree of error or when the outcome variable is merely a surrogate for an unobserved variable. For example, the survival time of cancer patients is frequently used as a surrogate variable for unobserved tumor subtypes. We apply our procedure to a series of simulated and real genetic data sets and show that it outperforms conventional methods and several competing approaches.

Basis Selection From Multiple Libraries

Jeffrey C. Sklar, CalPoly San Luis Obispo, Statistics Department; Junqing Wu, University of California, Santa Barbara, Statistics and Applied Probability Department; ✦Wendy Meiring, Department of Statistics and Applied Probability, University of California, Santa Barbara, Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110 USA, *meiring@pstat.ucsb.edu*; Yuedong Wang, University of California at Santa Barbara

Key Words: model selection, function estimation

We present recent results on a method for estimating complex functions by linear combinations of basis functions selected adaptively from different classes of basis functions called libraries. Libraries are chosen to model various features of a function such as change points and oscillations. Data-driven estimates of model complexities based on the generalized degrees of freedom are used to correct bias incurred by adaptive model selection. The proposed method is general in the sense that it can be applied to any generic libraries including spline and wavelet bases. Simulations and real data sets will be used for illustration.

470 Nonparametric Distance Measures and Tests ●

Section on Nonparametric Statistics, International Indian Statistical Association

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

A Generalized Multivariate Kurtosis Ordering And Its Applications

✦Jin Wang, Northern Arizona University, Department of Mathematics and Statistics, Flagstaff, AZ 86011-5717, *jin.wang@nau.edu*

Key Words: Kurtosis, Ordering, Depth function, Spread functional, Multivariate quantile function

It has been commonly admitted that the meaning of a descriptive feature of distributions is given by an ordering and that the measures for this feature are meaningful only if they preserve the ordering. However, while many multivariate kurtosis measures have been introduced, multivariate kurtosis orderings have received relatively little investigation. In this paper, we propose and study a generalized multivariate kurtosis ordering. Under some conditions, this ordering is affine invariant and determines elliptically symmetric distributions within affine equivalence. Some special cases of the generalized ordering provide the kurtosis orderings for various existing multivariate kurtosis measures. Those kurtosis orderings are applied to explore the relationships of the multivariate kurtosis measures. Some other applications of the generalized multivariate kurtosis ordering are also given.

Quantile Based Distances Among Random Variables

✦Reza Hosseini, Simon Fraser University, 8888 University drive, Burnaby, BC V5A 1S6 Canada, *reza1317@gmail.com*

Key Words: divergence measure, quantiles, distributions, invariant

We introduce measures of distance among random variables that are based on quantiles. We show that such measures are invariant under continuous strictly monotonic transformations. We investigate their relation to Kolmogorov-Smirnoff distance. A class of these distances are defined in such a way that while comparing two distributions more importance (weight) is given to different quantiles (for example to the tails). We use this distance measures to show optimal overall fits to a random variable are not necessarily optimal in the tails (and hence not appropriate to study extremes).

Incorporating Distance Informaiton In Comparing Categorical Data

✦Hao Chen, Stanford University, Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford University, Stanford, CA 94305, *haochen@stanford.edu*; Nancy Zhang, Stanford University

Key Words: compare categorical data, distance, high-dimensional, minimum spanning tree

When testing the equivalence of two multinomial distributions, the existing tests, such as Pearson's Chi-square test and the likelihood ratio test, do not work well when the number of observations is proportional to the number of categories. The χ^2 approximation to the existing tests with standard degree of freedom calculations is inaccurate, and the tests have low power even using permutation tests under this situation. To improve the power of the test, we use the implicit information of similarity between categories, which is usually naturally embedded in the data. We construct a graph on the observations based on several criteria - minimum spanning tree, cross-match and nearest neighbor - and count the edges connecting points from two different groups. An averaged statistic is used to deal with ties in the distance matrix due to the intrinsic property of categorical data. Simulation results show that the test based on minimum spanning trees works best because it makes fully use of the distance information. A Gaussian approximation to the permutation distribution of this test is also derived.

Nonparametric Tests For Equality Of Distributions Based On Precedence Probabilities

✦Rajarshi Dey, Kansas State University, 11 Dickens Hall, Department of Statistics, Manhattan, KS 66506, *rajarshi@ksu.edu*; Paul Nelson, Kansas State University

Key Words: Nonparametric methods, Precedence Probability, ROC, K-sample Problem

Let X(i) be a random variable with continuous distribution function F(i) ; i=1,2,...,K. Using precedence probabilities, we propose and explore a class of nonparametric tests based on independent random samples for the null hypothesis that all distribution functions are equal. One of these tests, designed to find any type of difference among the distributions, performs well in cases where the distributions differ only in location, only in scale or both in location and scale. Another member of this class is effective in testing for the volume under the Receiver Operating Characteristic surface , which, for K = 2 , is widely used in medical studies as a summary index of diagnostic accuracy. Comparisons to the performance of standard tests for these problems will be presented to show the proposed tests as a viable alternative to the standard procedures.

Model Specification Tests for Discretely Observed Semi-Martingales with Micro-Structure Noise

✦Zhi Liu, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, 999077 Hong Kong SAR, China, *liuzhi@ust.hk*

Key Words: Model specification, semi-martingales, micro-structure noise, power variations, high frequency data

To model a dynamic system, a financial security for instance, whether the underlying process jumps or not makes the model totally different. On the one hand, the imperfections of trading processes induce noisy observations, it is known as micro-structure noise. On the other hand, existing methods to test the jumps or diffusion are developed for the discretely observed data without noise. In this paper, we propose two procedures to test the presence of jumps and the presence of diffusion with noisy high frequency data. The proposed tests combine pre-averaging technique, two time scales method and threshold approach. Asymptotically, the test statistic is close to a known number under the null hypotheses, and it tends to other number otherwise for both tests. The theoretical asymptotic results are established. Simulation study confirms that the proposed procedures can successfully test jumps and diffusion as in the case without micro-structure noise. Finally, we implement the procedure to real high frequency financial data.

Outlier Detection For Multivariate Skew-Normal Data: A Comparative Study

✦Yinaze Herve Dovoedo, The University of Alabama, 302 Reed St. Apt 25D, Tuscaloosa, AL 35401 USA, *yhdovoedo@crimson.ua.edu*; Subhabrata Chakraborti, The University of Alabama

Key Words: outlier detection, data depth, outlyingness, skew-normal distribution

Detecting multivariate outliers is a challenging problem. It is common to use robust distances for that purpose. A more general way consists of using robust depth functions or equivalently through the corresponding "outlyingness" functions. Most outlier detection studies proposed in the literature assume that the underlying distribution of the data is normal. This paper deals with the case of multivariate skewed data, specifically, data following the multivariate skew-normal (Azzalini and Dalla Valle 1996) distribution, as considered in Hubert and Van der Veeken (2008). We compare the outlier detection capabilities of four robust outlier identifiers (through the corresponding outlyingness functions) in a simulation study. The simulation study explores two outlier occurring scenarios: "the cluster" and "the radial". Conclusions and recommendations are offered for each scenario.

Usage Of Distribution Extents As Cut-Off Parameters And Predictive Variables With Application To Web Analytics

✦Alex Zolotovitski, Microsoft Corp, 1 MSFT Way, Redmond, WA 98052, *alexzol@microsoft.com*

Key Words: distribution extents, web analytics

Distribution extents of order k for a sample {x1, x2,.. xn} of a non-negative stochastic variable X are defined as E_k = (sum p_i^k)^(1/(1-k)) if k != 1 , E_1 = exp(-sum(p_i ln(p_i))) if k = 1 , where p_i = x_i / sum(x_i) , and are useful measures of "a number of large values" in the sample. They were introduced by L.L. Campbell in 1964, and are generalization of inverse Herfindahl-Hirschman Index (HHI), a commonly accepted measure of market concentration in economics, and Simpson's diversity index used in ecology, and are closely related to Shannon-Wiener Index and the RÈnyi entropy and divergence. In this work we describe general properties of E_k and demonstrate advantages to use it in analysis of a web advertisement network, where actors are advertisers, publishers, and users, for two purposes: 1) as natural cut off parameters to present the network as a graph to visualize the network and to use graph theory methods and 2) as independent variables in predictive modeling.

471 Applications in the Physical and Life Sciences ■●

Section on Physical and Engineering Sciences, Section on Quality and Productivity, Section on Statistical Consulting, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Estimating Power Outage Duration In The Event Of Hurricane Land-Falls In The U.S.

✦Roshanak Nateghi, JHU, 4205 FALLS RD, APT12, BALTIMORE, MD 21211, *rnategh1@jhu.edu*; Seth David Guikema, JHU; Steven Quiring, Texas A&M

Key Words: power outage duration, hurricanes, data mining, power systems, Random forest

Each year, hurricanes cause widespread power outages in the United States that adversely affect industries and millions of customers, leading to large economic losses. Being able to model power outage durations accurately can facilitate more efficient coordination of recovery efforts and, as a result, help in minimizing the restoration periods. The objective of this paper is to develop statistical models to accurately estimate power outage duration prior to hurricane land-falls, using a wide range of covariates such as the hurricane's estimated wind speed and central pressure, the number of transformers, the number of customers, the number of poles, and a variety of geographic and development variables in the region of interest. Our models were developed using the data available from the outages caused by Hurricane Ivan in 2004. The results of this model were validated against the outage durations associated with Hurricane Katrina and Hurricane Dennis. Regression analysis together with data mining techniques were used to develop our statistical models, and we compared the models based primarily on their predictive accuracy.

Uncertainty Quantification In Building Energy Study - The Case Of Local Wind Speed

✦Huizhi Xie, Georgia Institute of Technology, 765 Ferst Drive NW, Atlanta, GA 30332, *hxie3@gatech.edu*; Yeonsook Heo, Georgia Institute of Technology; Godfried Augenbroe, Georgia Institute of Technology; C. F. Jeff Wu, School of Industrial and Systems Engineering, Georgia Institute of Technology

Key Words: Building Energy, Computer Experiments, Simulation, Uncertainty Quantification

EnergyPlus is a computer code widely used in simulating building energy consumption. The calculation of local wind speed in EnergyPlus, as an intermediate step of building energy consumption prediction, inherits uncertainty from the microclimate conditions caused by varying urban surroundings. This study aims to quantify the uncertainty of local wind speed calculation through the uncertainty quantification of wind speed profile exponent and boundary layer thickness. For a given terrain, different urban surroundings are generated using space filling

design coupled with random sampling. The outcomes of EnergyPlus are compared with those of a higher fidelity model (meso model) for each surrounding. A predefined measure of discrepancy is minimized to obtain a set of point estimates of the exponent and boundary layer thickness for each surrounding. From these estimates, the distribution of the two parameters can be estimated using smoothing techniques. The distribution of local wind speed induced from that of the two parameters will be utilized for further calculation in EnergyPlus. This is one of the first uncertainty studies in building energy.

Analysis Of Sonic Synthetic Profile In Oil Wells Through The Holt-Winters Models

✦Stephanie Russo Fabris, Federal University of Sergipe, Aracaju - SE, Brazil, *sterussofabris@hotmail.com*; Edmara Thais Menezes, Federal University of Sergipe; Maria Emilia Camargo, Universidade de Santa Cruz do Sul; Vitor Hugo Simon, PETROBRAS; Suzana Leit„o Russo, Federal University of Sergipe

Key Words: Time series, Holt-Winters, Sonic Profile

The sonic profile is used, mainly in exploration wells, which receive greater investment in data acquisition because these wells serve as a reference for analyzing the petroleum potential of the area. To be considered for exploitation wells are usually not racing sonic tools for economic reasons, especially in times of low oil price. Since it is not economically feasible to purchase a sonic profile, the vast majority of older wells, mainly due to the cessation of oil production, one must resort to other means to solve the problem. One way is to use statistical methods that generate a sonic profile, said synthetic, to establish estimates of seismic velocities of rocks along the wells. Thus, this study aims to determine a Holt-Winter model that generates synthetic sonic profiles in oil wells drilled in the geologic units belonging to the Sergipe-Alagoas sedimentary basin and that are within the required quality standards.

Using The Block Minima Method To Estimate Pipeline Thickness

✦Shiyao Liu, Iowa State University, 3426 Orion Dr. Apt 210, Ames, IA 50010, *lsy641@iastate.edu*; William Q. Meeker, Iowa State University

Key Words: Block size, Block minima method, Domain of attraction, Gumbel distribution, GEV distribution, Simulation

The block minimum method, based on extreme value theory, can be used to estimate the minimum thickness along a pipeline. Intuition suggests that knowledge about the domain of attraction of the thickness distribution should provide an important inferential advantage in that it should be possible to use the simpler appropriate two-parameter extreme value distribution (e. g., the Gumbel distribution) instead of the three-parameter generalized extreme value distribution to fit the data from the block-minimum method. An important question is how to choose the block size. We designed and conducted an extensive simulation to explore the impact of block size in the block-minima extreme value theory method for estimating small quantiles of a distribution of minima. Three parent distributions from the Gumbel minima domain of attraction were used: normal, the minimum Gumbel (or type 1 smallest extreme value or SEV) and the maximum Gumbel (type 1 largest extreme value or LEV). This paper summarizes the results of the simulation study to provide insight into the choice of the block size and extreme value distribution in the block minima method for modeling the pipeline thickness.

Application Of The Control Charts For The Evaluation Of Milk Quality

✦Suzana Leit„o Russo, Federal University of Sergipe, Aracaju - SE, 49035490 Brazil, *suzana.ufs@hotmail.com*; Evelyn Chagas, Federal University of Sergipe; Leda Santana, Federal University of Sergipe; Stephanie Russo Fabris, Federal University of Sergipe; Maria Emilia Camargo, Universidade de Santa Cruz do Sul

Key Words: Control Charts, SPC, Milk

This study aims to assess the milk quality supplied to LaticÌnios Industry Buril Ltda., located in the city Arau· (Brazil) during the period August 31, 2010 to September 19, 2010. For this we analyzed the variables density, freezing point and fat acidity using Shewhart control charts. It was obtained as a result for the Xbar control chart, the variable density values were of limits of the control. Regarding the control charts for the standard deviation, density, fat and acidity showed peaks, indicating the need to apply the controls charts revised to overcome this situation. Thus, it is essential that after the construction of control charts corrective actions are implemented during the production of milk, just when it looks at the failure, thus avoiding problems for the company and prevent further failures.

Comparison Of Methods For Setting Acceptance Criteria In A Multi-Staged Pharmaceutical Manufacturing Process

✦Richard Montes, Merck Research Laboratories, 2000 Galloping Hill Rd, K-15-2, Kenilworth, NJ 07033, *richard.montes@merck.com*; Jorge Quiroz, Merck Research Laboratories; David Christopher, Merck Research Laboratories

Key Words: acceptance criteria, multi-staged process, tolerance interval, risk minimization, variation transmission, autoregressive model

Pharmaceutical manufacturing processes consist of a series of stages (e.g., reaction, work-up, isolation) to generate the active pharmaceutical ingredient (API). Outputs at each stage need to be controlled within acceptance criteria to assure final drug product quality. In this paper, two methods based on tolerance interval to derive such acceptance criteria will be evaluated. The first method is SERIAL WORST-CASE, an industry risk minimization strategy, wherein input process parameters and materials of a stage are fixed at worst-case settings, the procedure serially repeated throughout next stages until the final output. The second method is the VARIATION TRANSMISSION USING AUTOREGRESSIVE MODEL wherein variation transmitted through ith stage is estimated accounting for the recursive structure of the errors at each stage. Computer simulations at varying extent of variation transmission and process variability are performed and the resulting confidence coefficients and average tolerance interval lengths in the two methods are compared. The second method performs better and is recommended for setting in-process and API streams in a multi-staged pharmaceutical manufacturing process.

Statistical Comparison Of Forensic Toolmarks

✦Amy B Hoeksema, Iowa State University, 1103 Pinon Dr Unit 1, Ames, IA 50014 US, *abh580@gmail.com*

Key Words: Forensics, Profilometry, Surface characterization, Matching, U-statistics, Correlated data

In forensics, fingerprints can be used to uniquely identify suspects in a crime. Similarly, a toolmark left at a crime scene can be used to identify the tool that was used. However, the current practice of identifying a match involves visual inspection of marks which can be a very subjective process. As a result of this subjectivity, declared matches are often successfully challenged in court, so law enforcement agencies are particularly interested in encouraging research in more objective approaches. We are working to develop a statistical algorithm that can be used to determine a match or non-match given a crime mark and a suspect tool. Our analysis is based on comparisons of profilometry data, essentially depth contours of a toolmark surface taken along a linear path. With sufficient development, such methods may lead to more defensible forensic analyses.

472 Advanced Topics in Regression ■●

Section on Physical and Engineering Sciences

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Measuring The Span Of The Visible Universe With Nonparametric Regression

✦Irina Udaltsova, UC Davis, Statistics Dept., One Shields Ave, Davis, CA 95616, *isudaltsova@ucdavis.edu*; Samuel Schmidt, UC Davis, Physics Dept.; Jiming Jiang, University of California, Davis

Key Words: astrostatisics, nonparametric, kernel regression

This talk presents an adaptive kernel regression technique for estimation in multiple dimensions. We apply the method to determine galaxy redshifts using photometry. The kernel regression is a nonparametric procedure for estimation of response variable based on locally weighted average of neighboring data points. The weights assume a distance-based kernel function with a pre-defined distance metric. Parallel to image processing and reconstruction ideas, a desirable property of the estimator is the recovery and noise removal of image. It seems appropriate that local image features should guide the estimation procedure. Orientation features in an image are summarized into image gradients, which re-define the distance metric according to the local dominant orientation. Adapted size and shape of the regression kernel give appropriate influence to local data points, resulting in preservation of image details. Such technique is proposed for use in multiple dimensions. Simulation and application to photometric data are also presented.

In Search Of Desirable Compounds

✦Adrijo Chakraborty, University of Georgia, 101 Cedar Street, Athens, GA 30602-7952, *adrijoc@uga.edu*; Kjell Johnson, Pfizer, Inc.; Abhyuday Mandal, University of Georgia

Key Words: Drug discovery, Desirability function, Partial least squares, Random forest regression, Unsupervised learning, Dimension reduction

In drug discovery, chemists often evaluate a compound's performance across a number of endpoints (efficacy, ADME properties, safety, etc.). A popular way to prioritize compounds is through desirability scoring across the endpoints of choice. In addition to desirability scores, chemists can measure or compute other compound descriptors that are thought to be related to the endpoints of interest. Applying the methods discussed here we identify some important descriptors and efficiently classify the compounds according to their desirability scores. Also, we explore ways to visualize this high-dimensional data that can effectively inform chemists' decisions in synthesizing new compounds

Designs For Discriminating Between Linear And Cubic Regression Models

✦Santanu Dutta, Department of Statistics, University of California, Riverside, Riverside, CA 92521, *sdutt001@ucr.edu*; Subir Ghosh, Department of Statistics, University of California,Riverside

Key Words: Model Discrimination, T-optimality, I and J criteria, Linear and Cubic Models, Optimal Design

We consider the identification and discrimination between two competing regression models, a linear and a cubic, with the response variable y and the explanatory variable x. The four distinct design points x1, x2, x3, and x4 are replicated n1, n2, n3, and n4 times respectively, satisfying n1 + n2 + n3 + n4 = n, n1 = n4 and n2 = n3. For a fixed value of n, we compare designs optimum with respect to two criteria, J (equivalent to T-optimality) and a proposed I. We obtain a class of designs that are better than the Dette-Titoff T-optimal (equivalently J-optimal) design (Dette and Titoff (2009)) under the criterion I in their setup where the quadratic coefficient is zero in the cubic model. However, the Dette-Titoff design is better than our class of designs under the criterion J. We also obtain optimal designs when the quadratic coefficient is not zero.

Selecting Linear Models Under The Bayesian Paradigm With Focus On Good Prediction Over A User-Specified Distribution On The Covariate Space

✦Adam Lee Pintar, National Institute of Standards and Technology, , *adam.pintar@nist.gov*; Christine Anderson-Cook, Los Alamos National Laboratory; Huaiqing Wu, Iowa State University

Key Words: Model Selection, Deviance Information Criterion, Posterior Probability, Bayesian Model Averaging, Correlated Variables

Model selection is an important part of building linear regression models in the Bayesian paradigm. If unimportant explanatory variables are included, posterior distributions will have inflated variance. If important explanatory variable are excluded, posterior distributions can miss their (unknown) target. Several model selection methodologies currently exist. For instance, one could choose the model with the smallest deviance information criterion, the model with the largest posterior probability, or the model whose terms all have posterior probability greater than 0.5. A common theme to all of these methodologies is that they consider only the observed data. We propose a model selection methodology that focuses on good prediction over a user-specified distribution on the covariate space. The methodology quantifies the prediction ability of all models under consideration at many covariate points sampled from the user-specified distribution. Then, a best model is identified by graphically comparing the distributions of prediction abilities. The methodology is illustrated via an example, and a simulation study highlighting its potential is presented.

Comparison Of Statistical Power For Testing Interactive Versus Quadratic Effects Using Measured Variable Versus Latent Variable Approaches

✦Yue Ma, University of South Florida, Office of Decision Support, BEH 245, 4202 E. Fowler Avenue, Tampa, FL 33620, *yuema@usf.edu*; Leona Aiken, Arizona State University; Stephen West, Arizona State University

Key Words: Statistical Power, Structural Equation Modeling, Latent Variable, OLS Regression, Interaction, Quadratic

The statistical power for detecting an interaction versus a quadratic effect using OLS regression for scale scores versus structural equation modeling (SEM) of latent variables was compared through simulation and analytic work. Models studied included two predictors and one nonlinear effect, either their interaction or a single quadratic term. OLS had greater power than SEM. OLS had attenuated parameter estimates, but smaller standard errors and MSEs relative to SEM for nonlinear terms. Comparisons equating the magnitude of the coefficients for the nonlinear terms showed greater power for the test of the quadratic term than the interaction term, which diminished as the interpredictor correlation increased. Comparisons equating the squared semi-partial correlations of the nonlinear terms showed equal power at low correlations, with the power for the interaction term exceeding that of the quadratic term as the interpredictor correlation increased. Results were interpreted in terms of (1) attenuation in parameter estimates for OLS, (2) negative bias in standard error estimates for SEM, and (3) differences between theoretical versus empirical standard errors for both OLS and SEM.

The Constrained Random Effect Models For Nanoparticle Synthesis In A Supercritical Co2 Process.

✦HIN KYEOL WOO, GEORGIA TECH, 2202 Druid Hills Reserve Dr NE, ATLANTA, GA 30329, *hinkyeol@gatech.edu*; Jye Chyi Lu, GEORGIA TECH

Key Words: Statistical modeling, Computer experiment, Generalized linear mixed model, Variance component

This research develops a two-part model for observations from nanoparticle fabrication experiments. Since there are certain combinations of process variables resulting to unproductive process outcomes, a logistic model is used to characterize such a process behavior. For the cases with productive outcomes a normal regression serves the second part of the model. Because the data are obtained from computer experiments, random-effects are included in both logistics and normal regression models to describe the potential spatial correlation among data. The likelihood function for this two-part model is complicated and thus the maximum likelihood estimation is intractable. This research studies approximation techniques based on Taylor series extension to simplify the likelihood. An algorithm is developed to find estimates for maximizing the approximated likelihood.

Mcmc Methods Using Temporary Mapping And Caching And Its Applications On Gaussian Process Regression

✦Chunyi Wang, University of Toronto, 100 St George Street, Room 6003, Toronto, ON M5S2M8 Canada, *chunyi@utstat.toronto.edu*; Radford M Neal, University of Toronto

Key Words: MCMC, Gaussian Process

We propose two general ideas for constructing efficient MCMC methods - temporarily mapping to a new space, which can be either larger or smaller than the original space, and caching the results of previous computations for future re-use. These two ideas can be combined to improve efficiency for a wide range of problems, examples include: problems where probabilities can be quickly recomputed when only some 'fast' variables are changed; effectively adapt tuning parameters without changing the Markov Chain transition (effectively adapting but not really adapting). We demonstrate the effectiveness of these methods by its applications on the Gaussian Progress Regression.

473 Section on Survey Research Methods - Imputation and Missing Data

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Microdata Imputation and Macrodata Implications: Evidence from the Ifo Business Survey

✦Christian Seiler, Ifo Institute for Economic Research, Poschingerstraße 5, Munich, 81679 Germany, *seiler@ifo.de*

Key Words: Business survey, Longitudinal data, micro data, Imputation, Nonresponse, macro data

Surveys are commonly affected by nonresponding units which can produce biases if these missing values can not be regarded as missing at random (MAR). As many papers examine the effect of nonresponse in individual or household surveys, only less is done in case of business surveys. This paper analyses the missing values in the Ifo Business Survey, which most prominent result is the Ifo Business Climate Index, a leading indicator for the businss cycle development in Germany. Covering the period from 1994 to 2009 (with about 7,000 responding firms per month) a considerable number of observations is available to compare various imputation approaches for longitudinal data to evaluate the method which reflects the underlying latent data generating process best. After this, the imputed data sets will be aggregated as usual and compared with the original indices to evaluate their implications to the macro level. We analyse the indicators for different aggregation levels and show that the bias is minimal.

Look Again: Editing and Imputation of SCF Panel Data

✦Arthur B Kennickell, Federal Reserve Board, Mail Stop 153, 20th and C Streets NW, Washington, DC 20551, *Arthur.Kennickell@frb.gov*

Key Words: Editing, Imputation, Panel data

In 2009, a re-interview with participants in the 2007 Survey of Consunmer Finances (SCF) was undertaken to provide information on the effects of the financial crisis on households. The panel questionnaire was designed to maximize comparability with the earlier data. The subject matter of the survey, wealth and related issues, is often considered sensitive or conceptually difficult. Consequently, editing and imputation of the data are very important considerations. Although the baseline data had already been edited and imputed cross-sectionally, they were re-edited along with the new panel data. Similarly, the data for both waves of the survey were imputed jointly. This paper has two goals: to examine the importance of the re-editing of the baseline data and to guage the effects of the joint imputation of data from the two waves.

Generalized Estimating Equation Model For Binary Outcomes With Covariates Missing By Design

✦Qixuan Chen, Columbia University, 722 West 168th Street, R652, New York, NY 10032, *qc2138@columbia.edu*; Myunghee C. Paik, Columbia University

Key Words: dioxins, estimating equations, jackknife estimator, missing data

We propose a method to handle monotone missing covariates in a generalized estimating equation (GEE) model for correlated binary outcomes, when the covariates are missing by design. The regression coefficients are obtained by solving an aggregate unbiased estimating function, and the variance of the regression coefficients is estimated using the one-step jackknife estimator (Lipsitz et al. 1994). The advantages of the proposed new method over the complete cases analysis and the inverse probability weighted estimating equation are demonstrated by simulation studies. The new method is used to study whether concentration of dioxin congeners in house perimeter soil is an important predictor for having a high concentration of dioxin congeners in household dust, where structurally similar and correlated congeners are present and a limited numbers of soil samples are measured.

Approaches For Handling Missing Sexual Identity Data

✦HarmoniJoie Noel, National Center for Health Statistics, Hyattsville, MD , *hnoel@cdc.gov*; Taylor Lewis, National Center for Health Statistics; Aaron Maitland, National Center for Health Statistics; Heather Ridolfo, National Center for Health Statistics

Key Words: missing data, imputation, sexual identity

Health disparities research in the Lesbian, Gay, Bisexual, and Transgender (LGBT) community has received national attention as part of the new goals for Healthy People 2020. Using data from the National Survey of Family Growth Cycle 6 (2002) and the 2002-2008 National Health and Nutrition Examination Survey we found that less educated and Hispanic respondents were systematically missing data on sexual identity measures. Cognitive testing results showed that less educated and Hispanic respondents had comprehension issues with some of the sexual identity terms which could help explain their higher rates of missing data on the sexual identity measures. Systematic missingness on sexual identity measures could lead to biased univariate estimates and regression analyses of the relationship between sexual identity and health outcomes such as access to health insurance or disease prevalence. We will compare alternative approaches for handling missing data including casewise deletion, single and multiple imputation, and maximum likelihood estimation to evaluate the sensitivity of univariate and multivariate estimates to these different approaches.

Imputation Of Categorical Data With Small Sub Categories: Assessing The Results Of Imputing Race In The 2007 - 2008 National Ambulatory Medical Care Survey (Namcs)

✦Elizabeth Goldberg, National Center for Health Statistics, 3311 Toledo Road, #3310, Hyattsville, MD 20782, *inw6@cdc.gov*

Key Words: imputation, categorical data, IVEware

The National Ambulatory Care Survey is a national probability sample survey of visits to nonfederal office-based physicians in the United States. Of particular interest to many researchers in health care policy is the effect of race on health care usage and outcomes, but race was missing at relatively high rates, peaking in 2008 at 33%. Evaluating the effectiveness and utility of the results of such imputations in categorical data, especially data with small sub categories, can be challenging. The evaluation of race imputation in NAMCS is a case study in how to evaluate categorical imputation. Graphical methods were used to compare imputed vs. non-imputed data. In addition, race was set to missing for a subsample of known cases in order to evaluate whether the model correctly imputed these known cases. The differences in the accuracy of the model were interesting because some very small categories were imputed with far more accuracy than some larger sub groups. This paper discusses the methodology and the findings for the research.

Evaluation Of Improvements To The Imputation Routine For Health Insurance In The Cps Asec

✦Miche Boudreaux, University of Minnesota, 2221 University Ave. SE, Suite 345, Minneapolis, MN 55414, *boudr019@umn.edu*; Joanna Turner, University of Minnesota; Michael Davern, NORC at the University of Chicago

Key Words: CPS ASEC, health insurance coverage, imputation

The Annual Social and Economic Supplement of the Current Population Survey (CPS ASEC) is a widely used source of data on health insurance coverage. About 10 percent of respondents do not answer any of the ASEC supplement questions and the entire supplement for these respondents is imputed. Davern et al. (2007) identified problems with the imputation of health insurance variables in the CPS ASEC. They found the full supplement imputations cases were less likely to have private coverage and more likely to be uninsured. This discrepancy may partially be caused by a misspecification of the hot-deck routine. Only members of a policy holder's nuclear family can be covered as dependents. This family restriction contrasts with the instrument itself. We collaborated with the U.S. Census Bureau and evaluated possible improvements to the imputation routine to fix the identified problems. The study team recommended switching the order of the hot-deck allocation matrices with public coverage first and removing the nuclear family restriction to the private coverage allocation routine. Using 2009 CPS ASEC data, making these modifications reduces the uninsured rate from 17.3% to 16.8%.

Erosion Analysis With Usle And Rusle2

✦Yang Li, Iowa State University, 2505 Aspen Rd Unit 3, Ames, IA 50010, *yangli@iastate.edu*; Zhengyuan Zhu, Iowa State University

Key Words: USLE, RUSLE2, erosion, random forest, MARS

Universal Soil Loss Equation is a model that predicts the long term average annual rate of soil erosion caused by rainfall. It will be replaced by the more sophisticated RUSLE2 (Revised USLE 2) model in 2006. Data for computing both USLE soil loss estimates and RUSLE2 estimates are collected between 2002 and 2006 for validation. To estimate long term soil erosion trend and maintain consistency as an input to other models, it is necessary to impute USLE estimates after 2006 and RUSLE2 estimates before 2002. We first present statistical models to predict soil loss estimates from one soil loss model using the variables from the other one. With the aid of cross validation, we assess three different fitting methods: simple linear regression, multiple linear regression (MLR), and multivariate adaptive regression splines (MARS). Next we investigate imputation of USLE estimates shortly after 2006 and RUSLE2 estimates shortly before 2002 using nearest neighbor, classification and regression tree, and random forest. Reliable imputation is achieved by exploiting the relationship among the changes between neighboring years.

474 Analysing Data From Genome and Micro Arrays

Biopharmaceutical Section

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Comparisons Of Confirmation Rates Among Three Hit Selection Methods In A High-Throughput Screening Assay

✦Minya Pu, University of California, San Diego, , *mpu@ucsd.edu*; Karen Messer, Moores UCSD Cancer Center; Howard Cottam, University of California, San Diego ; Tomoko Hayashi, University of California, San Diego; Maripat Corr, University of California, San Diego ; Dennis Carson, University of California, San Diego

Key Words: High-throughput screening (HTS), Cancer vaccine, Computational chemistry, Murcko classes, Enriched structural family, Confirmation rate

A high-throughput screening (HTS) assay was performed to identify small molecules that stimulate innate immune receptors in cancer vaccine adjuvant design; ~160,000 compounds underwent the primary screen. Computational chemistry methods along with the naÔve 'top X' approach were used to screen active substances ('hits') in silico. Two enriched structural family approaches, Daylight clusters and Murcko classes, were first applied and then the top X approach was used for the remaining data. A total of 2033 hits were carried forward to the confirmation screen. A hit was considered 'confirmed' if its confirmatory activity was above a threshold so that ~600 compounds would be obtained for further cosideration. Confirmation rates(CR) were 31.7%, 42.4%, or 20.3% for the three approaches, respectively, in the order as stated above. The unbiased top X approach was also considered, that is, to call a hit from all the data (vs. remaining data). All the 1409 compounds that had activity values available from both the screens were included in the re-calculation. The CR was 32.9% for the unbiased top X approach, 47.1% for the Daylight clusters approach and 57.1% for the Murcko classes approach.

Evaluating Surrogate Variables For Improving Microarray Multiple Testing Inference

✦Jared Lunceford, Merck Research Laboratories, 4087 Foothill Dr, Provo, UT 84604, *jared_lunceford@merck.com*; Guang Chen, Merck Research Laboratories; Peter Hu, Merck Research Laboratories; Devan V Mehrotra, Merck Research Laboratories

Key Words: surrogate variables, multiple testing, testing dependence, microarray

The use of surrogate variables has been proposed as a means to capture, for a given observed set of data, sources driving the dependency structure among high-dimensional sets of features and remove the effects of those sources and their potential negative impact on simultaneous inference. In this article we illustrate the potential effects of latent variables on testing dependence and the resulting impact on multiple inference, we briefly review the method of surrogate variable analysis proposed by Leek and Storey (2008), and assess that method via simulations intended to mimic the complexity of feature dependence observed in real-world microarray data. The method is also assessed via application to a recent Merck microarray data set. Both simulation and case study results indicate that surrogate variable analysis can offer a viable strategy for tackling the multiple testing dependence problem when the features follow a potentially complex correlation structure, yielding improvements in the variability of false positive rates and increases in power.

Batch Effect Removal Method For Microarrays: Batch Profile Estimation, Correction, And Scoring

✦Tzu-Ming Chu, SAS institute Inc., S4036, SAS Campus Dr., Cary, NC 27513, *tzu-ming.chu@sas.com*; Wenjun Bao, SAS Institute Inc.; Russell Thomas, Hamner Institute for Health Services; Russ Wolfinger, SAS Institute Inc

Key Words: batch correction, batch profile, batch scoring, normalization, cross validation

Batch effect increases the variation among expression data and, hence, lowers the statistical power for investigating biological effects. When the corresponding proportion of variation associated with the batch effect is high, it is desirable to remove the batch effect from data. A robust batch effect removal method should be easily applicable to new batches. Here we discusse a simple, but robust grouped-batch-profile (GBP) normalization method that includes three steps: batch profile estimation, correction, and scoring. Genes with similar expression patterns across batches are grouped. The method assumes the availability of control samples in each batch, and the corresponding batch profile of each group is estimated by an ANOVA. Batch correction and scoring are based on the estimated profiles. A mouse lung tumorigenicity data set is used to illustrate GBP normalization through cross validation on 84 predictive models. On average, cross-validated predictive accuracy increase significantly from 0.56 to 0.66 after GBP normalization.

Multiple Contrast Tests For Testing Dose-Response Relationship In A Pre-Clinical Microarray Experiment

✦Setia Pramana, I-Biostat, Universiteit Hasselt, Agoralaan - building D, Diepenbeek, International 3500 Belgium, *setia.pramana@uhasselt.be*; Dan Lin, I-Biostat, Universiteit Hasselt; Ziv Shkedy, I-BioStat, Hasselt University

Key Words: Microarray, Dose-response, Multiple contrast tests, Isotonic regression, Gene expression, Trend test

Dose-response studies are the fundamental study in pharmaceutical industries to investigate the dose dependency of the response. In recent years dose-response studies have been integrated with microarray technology where several microarrays are available for a sequence of increasing dose levels and the response is gene expression. To identify genes with expression levels that change with studied dose levels, several tests for homogeneity of the means against order restricted alternatives in a microarray setting are available: the likelihood ratio test, Williams, Marcus, M and the modified M. We propose the implementation of Multiple Contrast Tests to these tests. The idea is that the order restricted alternative hypotheses can be decomposed in several elementary alternatives with particular patterns of equalities and inequalities. These contrasts are tested and the maximum contrast test is used. The inference is using its asymptotic distribution, permutations and the Significance Analysis of Microarrays. The power of these tests is compared using a simulation study to investigate the performance of these procedures. A dose-response study with 6 dose levels and 11,562 genes is used.

Correction For Population Stratification In Genome-Wide Association Studies

✦Li Liu, Sanofi-Aventis, 200 crossing blvd, bridgewater, NJ 08807, *li.liu@sanofi-aventis.com*; Donghui Zhang, Sanofi-Aventis

Key Words: population stratification, genome wide association studies

Genome-wide association studies can be useful for us to understand the diseases of interests and possibly identify the most proper treatments for each individual. Correcting for population stratification is very important in genome-wide association studies since population stratification can cause spurious associations. Currently, principal component analysis is widely used to deal with population stratification and has proven to be useful. However, it treats SNP data as continuous and cannot handle missing values. We propose to use modified singular value decomposition approach that treats the SNP data as ordinal, which is the true nature of the data, and at the same time, the modified approach will be able to handle missing values and outliers. We will demonstrate our method using both simulations and real datasets.

nPARS: A Comprehensive Search Algorithm for Constructing Bayesian Networks Using Large-Scale Genomic Data

✦Yen-Yi Ho, Johns Hopkins University, 100 W. University Pkwy, Apt 4H, Baltimore, MD 21210 United States, *yho@jhmi.edu*; Leslie Cope, John Hopkins University; Giovanni Parmigiani, Harvard/Dana-Farber

Key Words: Chemotherapy resistance, bayesian networks, large-scale genomic data, search algorithms, eQTL study

Background: Several studies have shown that the resistance to docetaxel and 5-FU in human cancer cell are significantly inheritable, but little is known about the underlying regulatory networks for the chemoresistance responses. Conventional gene association studies have reported on genomic location that could have effects on these differential treatment responses. However, much of the etiologic pathway mechanisms still remain largely unknown based on these study results. In this paper, we consider an extension of the traditional gene association study design by integrating information about genomic DNA variations and gene expression abundance and infer genetic networks that are associated with cellular responses to docetaxel and 5-FU. Results: We propose a comprehensive search approach named network partition reassembly search (nPARS) algorithm to construct Bayesian networks using large-scale genomic data. Based on our simulation results, we suggest that the proposed nPARS algorithm achieved better performance than commonly used search algorithms in terms of computational efficacy. We also demonstrate the implementation of nPARS algorithm in an experimental data set.

475 Issues with Predictive Power and Probability of Success ■

Biopharmaceutical Section

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Performance Of Some Proof Of Concept Decision Rules For Multiple Response Variables In A Multi-Dose Clinical Trial

✦Yiyun Tang, Pfizer Inv, 10777 Science Center Driver, San Diego, 92121 USA, *Yiyun.Tang@Pfizer.com*

Key Words: Multiple Response Variables, clinical trial design, type I error, Multiple Dose, Proof of concept

The proof of concept clinical trial often requires evaluation of multiple response variables for two or more doses of a test medicine. An example is a proof of concept study by a pharmaceutical company for the treatment of a psychiatric disorder. This paper describes statistical methods for a study design with multiple doses and balanced sample size, and it proposes reasonable statistical properties decision rules. Simulation studies are conducted to evaluate alternative decision rules with different degrees of flexibility. Discussion addresses the influence of correlations among the multiple response variables on the type I error, and the influence of the pattern of treatment differences on statistical power. A rule which requires a relatively strong result for at least one response variable and at least moderately favorable trends for all but one of the others is found to provide good statistical power with adequate control of type I error.

Predictive Power to Assist Phase III Go/No-Go Decision

✦Shengyan Hong, MedImmune, Inc, One MedImmune Way, Gaithersburg, MD 20878, *hongs@medimmune.com*; Li Shi, MedImmune, Inc

Key Words: Predictive power, phase 3 go/no go decision, hazard ratio, progression-free survival, overall survival

One of the most critical decision points during drug development is to make a phase 3 go/no go decision after a phase 2 proof of concept (POC) trial is conducted. It is particularly challenging in oncology drug development where oftentimes the primary endpoint for phase 3 trial is overall survival (OS) but the phase 2 POC trial is powered only for an early endpoint, typically progression-free survival (PFS), whose relationship to OS is often unclear. We propose the use of predictive power to assist the phase 3 go/no go decision by evaluating the strength of actual observed phase 2 efficacy outcomes in terms of how likely it will predict the chance of OS success in subsequent phase 3 trial. The formula is provided for calculation of predictive power based on either observed PFS outcome only, or observed OS outcome only, or both. An example is provided to compare these three predictive powers. It shows that when there is little prior information about OS and PFS, the observed PFS outcome has low predictability itself and also has limited added value to the predictability based on the observed OS outcome only.

A Multivariate Performance Function For Comparing Phase I Dose Escalation Designs

◆Dominik Heinzmann, F.Hoffmann-La Roche Ltd., Malzgasse 30, Basel, 4070 Switzerland, *dominik.heinzmann@roche.com*

Key Words: dose escalation designs, conditional probability of success, comparison

In Phase I dose escalation trials in Oncology, different designs as the traditional 3+3 method, the continual reassessment method or even more sophisticated methods jointly modeling safety and efficacy (e.g. biomarker) can be applied to find an appropriate dose for an investigational drug. These designs can be compared with respect to measures such as sample size, trial duration, accuracy and precision of the estimated "optimal" dose. In this paper, a unified measure is introduced enabling assessment of the overall performance of dose escalation designs. This holistic function enables comparison of the designs. It incorporates the conditional probability of success in a sequential Phase II trial given the selected dose in the Phase I to adjust sample size for estimation accuracy and precision. Simulation studies indicate that this performance function is a valuable tool for selecting an appropriate dose escalation method for a new drug. In particular, the function can be useful to evaluate if biomarker information should be included into the dose escalation framework.

Probability Of Statistical Significance In K Of M Future Trials Conditional On Current Observed Data

◆Jenny J. Zhang, U.S. Food and Drug Administration, 10903 New Hampshire Ave., Silver Spring, MD 20993, *jenny.zhang@fda.hhs.gov*

Key Words: Statistical power, Probability of statistical significance, Phase III trial design

In a separate paper to be presented, the probability of statistical significance (PoS) of a future trial is formulated with an expected statistical power over a fiducial distribution determined according to the likelihood function of treatment effect conditional on current observed data. It is also proven that the outcomes of multiple future trials conditional on the observed data are correlated due to the common condition. Here, we extend those results to a clinical development setting where m phase III trials are planned with an anticipation of at least k trials (k < = m) showing statistical significance to obtain a regulatory approval. The PoS of at least k of m trials can be formulated as the summation of expected statistical powers, which can be evaluated using the joint distribution of the future outcomes. Statistical proof and computational procedures with illustrative examples of the PoS of at least k of m future trials are presented in this paper.

Estimating The Probability Of Statistical Significance In Two Future Trials Given Current Observed Data

◆Jianliang Zhang, MedImmune, LLC, One Medimmune Way, Gaithersburg, MD 20878, *zhangj@medimmune.com*; Jenny J. Zhang, U.S. Food and Drug Administration

Key Words: Probabilility of success, statistical power, probability of simultaneous successes

At a clinical drug development stage (e.g., the end of phase II), the sponsor needs to evaluate the probability of achieving statistical significance (PoS) in the future phase III trial(s). The usual statistical power of the phase III could serve as the PoS if and only if the power is evaluated for the true effect size, which is unknown. However, a likelihood function can be determined according to the observed phase II data, and the PoS can be appropriately evaluated as an expected power based on that likelihood function. While two trials are usually required to be statistical significant in a phase III program, the future outcomes of the trials conditional on the same observed data are proven to not be statistical independent. Thus, the probability of simultaneous statistical significance (PoSS) of the two trials should be evaluated as the expected power over the joint distribution of the future outcomes, not a simple multiplication of the PoS of individual trials. Statistical proof and numerical illustrations for determining the PoSS are presented in the paper.

476 Issues in Vaccines ■

Biopharmaceutical Section, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

An Alternative To Ancova When Standard Assumptions Are Untenable

◆J. Brooke Marshall, Merck Research Laboratories, 2 N Ridge Ave Apt 2F, Ambler, PA 19002, *jemarsh2@vt.edu*; Devan V Mehrotra, Merck Research Laboratories

Key Words: Analysis of Covariance, Discretization, Minimum Risk Weights, SP-EFFIC method, Stratification

In a simple analysis of covariance (ANCOVA) analysis for a randomized clinical trial, it is assumed that the response variable (Y) is linearly related to the baseline covariate (X), and that the variance of Y is constant across all levels of X. A violation of either assumption can result in a suboptimal analysis. We propose an alternate approach in which the continuous variable X is discretized to form three categorical strata, and a stratified analysis using "minimum risk" stratum weights (Meh-

rotra and Railkar, 2000) is then used for estimation and inference. The operating characteristics of our proposed approach will be contrasted with those of the usual ANCOVA and the SP-EFFIC method of Tsiatis et al. (Stat Med, 2007) using simulations. A vaccine clinical trial will be used for illustration.

Threshold Methods for Immunological Correlates of Protection

Andrew Dunning, Sanofi Pasteur; Fabrice Bailleux, Sanofi Pasteur; Li Qin, Fred Hutchinson Cancer Research Center; ✦Xuan Chen, Sanofi Pasteur, , *xuan.chen@sanofipasteur.com*; Kamal Desai, University of London

Key Words: vaccine, correlate of protection, correlate of immunity, protective threshold, immunological assay

Immunological correlates of protection, here we refer to the immunological responses which are associated with protection from disease, play a central role in vaccines research. Although in reality the relationship is likely continuous: increasing assay values corresponding to increased protection from disease, in practice the use of threshold protective levels is ubiquitous. Yet limited quantitative methods exist for inferring protective thresholds from data on immunological assay values and subsequent disease occurrence. Standard methods are often unsuitable, since in most study designs a large number of unprotected subjects do not develop disease, in addition to those protected. This research examines nine threshold models and methods in terms of their suitability for estimating threshold immunological correlates of protection, and develops tests and measures of discrimination and goodness-of-fit for evaluating thresholds so found. Results are illustrated using fifteen datasets.

Vaccine Efficacy In The Intention-To-Treat Population: When Is A Pooled Analysis A Fooled Analysis?

✦Oliver Bautista, Merck & Co. Inc., P.O. Box 1000, UG1CD-44, North Wales, PA 19454 U.S.A., *oliver_bautista@merck.com*

Key Words: Prophylactic vaccine efficacy, intention-to-treat analysis

A prophylactic vaccine is efficacious in preventing a target disease when administered to subjects who are unexposed to vectors causing the disease. In a paradigm where vaccination is administered without pre-screening for exposure to vectors causing the target disease, assessment of vaccine efficacy (VE) in the intention-to-treat (ITT) population is the preferred approach for assessing overall population benefit of vaccination. For a prophylactic vaccine that makes no claim of a therapeutic effect among the sub-population of subjects who have already been exposed to vectors causing the target disease, an ITT evaluation of VE based on a pooled analysis is misleading and should not be the standard approach for estimating prophylactic VE in the entire population. Illustrations using data on a clinical trial of a quadrivalent HPV vaccine will be provided.

Bayesian Variable Selection Methods Applied To Antibody Response To Influenza Vaccination In Healthy Elderly Adults

✦Kimberly Crimin, Vanderbilt University Medical Center, Department of Biostatistics, , *kimberly.crimin@vanderbilt.edu*; Yuwei Zhu, Vanderbilt University Medical Center, Department of Biostatistics; Keipp Talbot, Vanderbilt University Medical Center, Department of Medicine; Marie Griffin, Vanderbilt University Medical Center, Department of Preventative Medicine; Kathryn Edwards, Vanderbilt University Medical Center, Department of Pediatrics

Key Words: influenza vaccine, bayesian variable selection, BMA, elderly adults

There has been concern about whether the influenza vaccine offers sufficient protection in elderly since immune response in the elderly are known to be less vigorous than in younger adults. Declining immune function in older adults is a hallmark of aging which affects the ability of this vulnerable population to respond to vaccination against influenza. The age-associated changes in the immune system occur in different people at different rates and are compounded by chronic medical conditions, making healthy elderly adults a diverse population. Healthy elderly adults were recruited and vaccinated with seasonal influenza vaccine in fall. To determine the subset of demographic, clinical and immunogenicity variables that were useful in predicting the post-vaccination antibody response to influenza vaccine, several common Bayesian variable selection methods were applied to the data. To evaluate the predictive ability of the top models (models with the highest posterior probability) selected by the method, Bayesian Model Averaging was used. In this talk, we present the results from the various methods and discuss the different approaches for the motivated vaccine research.

Evaluation Of Immune Response (Gpelisa) As The Principal Surrogate Endpoint For Protection Of Herpes Zoster Afforded By Zostavax

✦Xiaoming Li, Merck Research Laboratories, 19454, *xiaoming_li@merck.com*; Xiaopeng Miao, Department of Biostatistics, Boston University; Ivan S.F. Chan, Merck Research Laboratories

Key Words: Surrogate Endpoint, Statistical Surrogacy, Principle Surrogacy, Vaccine Trial

Use of surrogate endpoints in clinical trials can shorten and hence make drug development more cost-effective. Validating a surrogate for a clinical endpoint is critical in this context. The most commonly used approach to validate a surrogate endpoint is to demonstrate 'statistical surrogacy', based on the four criteria by Prentice (1989). A more recently proposed approach is 'principle surrogacy', by Frangakis and Rubin (2002) and Gilbert and Hudgens (2008). In the herpes zoster (HZ) vaccine (ZOSTAVAX) efficacy trials, the primary endpoint is the incidence of HZ. The antibody response, as measured by gpELISA post-vaccination, has been established as an immune correlate of protection based on these efficacy trials, using the approach of statistical surrogate. In this work, the 'principle surrogacy' of gpELISA will be evaluated, together with its 'statistical surrogacy'.

Comparison Of Tlovr Algorithm Versus Snapshot Approach For Hiv Studies

Wei Zhang, Boehringer Ingelheim Pharmaceuticals, Inc.; Michael Pannucci, Boehringer Ingelheim Pharmaceuticals, Inc.; ✦Junhai Guo, Boehringer Ingelheim Pharmaceuticals, Inc., , *junhai.guo@boehringer-ingelheim.com*; David Hall, Boehringer Ingelheim Pharmaceuticals, Inc.

Key Words: TLOVR algorithm, SNAPSHOT approach, HIV studies, Phase III

The time to loss of virologic response algorithm proposed in the FDA guidance for industry on antriretroviral drugs using plasma HIV RNA measurements (2002) has been used for sponsors traditionally for submission of clinical trials in the indication of HIV-1. This algorithm takes into account a patient's logitudinal viral load up to the pre-defined time point by considering patterns of suppression and rebounding. Recently, FDA is evaluating if this algorithm can be simplified using the SNAPSHOT approach, i.e., defining a patient's virologic response status using only the viral load at the pre-defined time point within an allowed window (Smith et al, 2009 JSM presentation). We compare the differences between the TLOVR algorithm and SNAPSHOT approach and apply both methods to two phase III pivotal trials: VERxVE and TRANxITION. The results suggest that the simple SNAPSHOT approach is a reasonable alternative to the complicated TLOVR algorithm.

477 Panel Data Methods and Applications ■

Business and Economic Statistics Section

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Nonlinear Detrending In Panel Models To Estimate Macroeconomic Effects On Mortality

✦Zhen Wang, University of Michigan, 1085 South Universirty,439 West Hall, ANN ARBOR, MI 48109, *wzhen@umich.edu*; Edward L. Ionides, University of Michigan; Jose A. Tapia Granados, University of Michigan

Key Words: PANEL, HP, fixed-effect

Previous investigations of the effect of macroeconomic conditions on mortality often used fixed-effect panel models applied to mortality rates and unemployment rates in levels (or log-transformed). We show that this type of panel models often generates spatio-temporal patterns of the residuals and therefore violation of the regression model assumptions, that probably bias the effect estimates. Using panel regression models applied to differenced or Hodrick-Prescott detrended data, the high serial correlation of the observations is thoroughly removed and the patterns of the residuals disappear. This kind of panel models applied to previously detrended data of the US states confirm that mortality oscillates procyclically.

A Simple Panel Stationarity Test In The Presence Of Cross-Sectional Dependence With An Application

✦Kaddour Hadri, Queen's University Belfast, 25 University Square, Belfast, BT7 1NN UK, *k.hadri@qub.ac.uk*

Key Words: Panel data, stationarity, KPSS test, cross-sectional dependence, joint asymptotic, Prebish-Singer hypothesis

This paper develops a simple test for the null hypothesis of stationarity in heterogeneous panel data with cross-sectional dependence in the form of a common factor in the disturbance. We do not estimate the common factor but mop-up its effect by employing the same method as the one proposed in Pesaran (2007) in the unit root testing context. Our test is basically the same as the KPSS test but the regression is augmented by cross-sectional average of the observations. The limiting distribution under the null is shown to be a standard normal. The latter result is derived using the joint asymptotic limits where T and N?8 simultaneously (under the additional condition that N/T?0). We also extend our test to the more realistic case where the shocks are serially correlated. We use Monte Carlo simulations to examine the finite sample property of the panel augmented KPSS test. An application is also provided where the proposed panel stationary test is applied in the first step when testing the secular decline of real commodity prices known as the Prebish-Singer (PS) hypothesis.

Estimating Panel Cointegration Models With Global Stochastic Trends -- Heterogeneous Panel And Adjustment Of Error Structures

Qian Lin, Southwest University of Finance and Economics; ✦Jun Liu, Georgia Southern University, GA 30458, *jliu@georgiasouthern.edu*

Key Words: panel cointegration

This paper is concerned with the estimation of panel cointegration models with cross-sectional dependence generated by unobserved global stochastic trends. We extend the Continuously Updated and Fully Modified Estimators (CupFM, Bai et al. 2009) under the condition of cross sectional heterogeneity. To improve the precision of estimation, we propose to adjust the estimated error structure. The limiting distributions of the estimators are derived. The finite sample properties are studied by Monte Carlo simulations.

Extending Dynamic Factor Analysis To Panel Data: An Application On Benchmarking Index For Water Utilities

✦Nikolaos Zirogiannis, University of Massachusetts, Amherst, 404 Stockbridge Hall, 80 Campus Center Way, Department of Resource Economics, Amherst, MA 01003, *nzirogia@resecon.umass.edu*; Yorghos Tripodis, Boston University; Joe Moffitt, University of Massachusetts, Amherst; Alexander Danilenko, World Bank

Key Words: Dynamic Factor Analysis, Panel Data, EM algorithm, Leading indicator, Water utilities

We develop a dynamic factor analysis model that can be applied on panel data. While dynamic factor analysis is a widely used statistical technique, there is limited research regarding its application on panel data. In simple dynamic factor analysis, we have multiple variables of the same subject, such as a company or a country measured at each time period. In a more general panel data setting, we have several variables of multiple subjects, such as a company, nested within a higher level variable, such as a country, measured at each time period. This creates an extra element of complexity, since we need to account for between company variability, as well as within and between country variability. In this paper, we propose a simple solution to this problem by using the EM algorithm to estimate factor loadings for panel data. We explore different correlation structures for the within-country variability and propose a model selection methodology. The model is then applied on a dataset of various countries, each with multiple water utilities. The goal is to create a leading benchmarking index that will be able to predict the inability of a utility to cope with external stresses.

Dynamic Regression Models With Holiday Effects And Sarma Errors For Forecasting Short-Term Electricity Demand In Korea

✦Myung Suk Kim, SSME dept., Sogang Business School, Sogang University, South Korea, *myungkim@sogang.ac.kr*

Key Words: double SARMA model, dynamic regression model, forecasting, holiday day effects, short term electricity loads

Dynamic regression models are applied to the prediction of hourly electricity demand in Korea that contain multiple cycles and holiday effects. The suggested dynamic regression models include the interaction terms of time effects and holiday effects in the transfer function. The stochastic disturbance is modeled using the double seasonal autoregressive moving average (SARMA) mechanism. One day ahead forecasting performance of the suggested model is compared with the traditional double SARIMA models over the next one year horizon. Our empirical results indicate that the suggested model outperforms the benchmark models.

Spatially Correlated Unbalanced Longitudinal Energy Billing Data Analysis

✦Min Niu, University of Wisconsin, Madison, Department of Statistics, 1300 University Ave., Madison,, WI 53706, *niu@stat.wisc.edu*; Zhengjun Zhang, University of Wisconsin

Key Words: spatial panel data, energy bills

Residential energy bills are featured with unequally spaced longitudinal data which are collected by utilities. The collected data also demonstrate spatial correlation. In energy efficiency program evaluations, billing data based analysis is one of the major methods, yet it has received little discussion in the recent past. Motivated by this billing analysis, this study proposes a varying-coefficient multilevel mixed-effects model with spatial correlation introduced through the community level effects, and an auto-correlation structure within each household. In the proposed model, the efficiency of regression coefficient estimators depends on how accurately spatial and time-wise covariance parameters are estimated. The widely used restricted maximum likelihood method is not applicable to the proposed model as optimization is a problem. A two step estimation method is employed for estimating covariance matrices. The proposed estimators are consistent, which provide a feasible way to solve the optimization problem with satisfied precision. Sampling properties and finite-sample performance are presented. Real energy billing data is analyzed.

Modeling Hourly Day-Ahead Electricity Prices: A Functional Data Analysis Approach

✦Asitha Edirisinghe, Missouri University of Science & Technology, Department of Mathematics & Statistics, 400 Westh 12th Street, Rolla, MO 65409-0020, *eae938@mail.mst.edu*; V A Samaranayake, Missouri university of Science and Technology

Key Words: High Frequency Time Series, Price Forecasting, Electricity Market, Spline Functions, Irregular Seasonality, Volatility

Four years of hourly day-ahead electricity price data from the region supplied by the American Electric Power (AEP) company through the PJM Regional Transmission Organization (RTO) are modeled. The price data show several features, such as irregular seasonal behavior, weekly and daily cycles, as well as sensitivity to oil price fluctuations, making it unsuitable for modeling using techniques that assume stationarity. We adopt a structural time series approach to remove some of the non-stationary behavior in the daily aggregate series and model the residual hourly prices obtained by subtracting the predicted daily average from the hourly series. The series is then scaled to remove long-term fluctuations in the variance due to economic conditions and demand. The resulting series is then modeled using spline functions to estimate the daily cycles that change across summer, winter and shoulder months. A GARCH type model is also used to investigate the within-day conditional volatility that is observed in the price data. This case-study can be looked upon as providing a new paradigm for modeling high-frequency data with non-stationary and cyclical features that change over time.

478 High-dimensional Covariance Matrix Inference ●

IMS

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

High Dimensional Sparse And Low Rank Covariance Matrix Estimation Via Convex Minimization

✦Xi Luo, University of Pennsylvania, 3730 Walnut St, 400 JMHH, Philadelphia, PA 19104 United States, *xi.rossi.luo@gmail.com*; Tony Cai, University of Pennsylvania

Key Words: lasso, low rank, sparse, covariance matrix, factor model, random effect

We propose a new approach based on convex minimization to estimate a general class of covariance matrices that can be decomposed as a sum of low rank and sparse components. Many classical statistical models, such as factor models and random effect models, motivate this covariance structure. The resulting estimator is shown to recover exactly the rank and support of the low rank and sparse components respectively. The convergence rates under the spectral norm and the elementwise-infity norm are also presented. We propose iterative algorithms to solve the optimization criterion. The algorithm is shown to be within an order 1/t^2 neighborhood of the optimal after any finite t iterations. Numerical performance is illustrated using simulated data and stock portfolio selection on S&P 100. This is joint work with T. Tony Cai.

Test For The Equality Of Covariance Between The Two Components Of Random Vectors For High Dimensional Data

✦Jun Li, Department of Statistics, Iowa State University, Ames, IA 50011, *junli@iastate.edu*; Song Xi Chen, Iowa State University

Key Words: high dimension, covariance matrices

We test if the covariance between the two components of the high dimensional random vectors are the same for the two underlying population. This interest arises naturally in the gene-sets testing, where we are interested to know if any two segments of the gene-sets have the same covariance between two treatments. Therefore, except testing if two covariance matrices are equal to each other in the overall structure, we move further and include a test which can compare the equality of the covariance between the two components of random vectors.

Hypothesis Testing For High-Dimensional Data

✦Yingli Qin, Nanyang Technological University, 21 Nanyang Link, SPMS-MAS-05-02, Singapore, 637371 Singapore, *qinyl@ntu.edu.sg*

Key Words: high-dimensional, testing, manova, martingale, asymptotic normal

We propose a high-dimensional MANOVA test statistic, in the case where both the sample sizes and dimensionality grow to infinity. The asymptotic distribution of the test statistic has been shown to be asymptotic normal under general conditions using martingale central limit theorem. Simulation results will be provided to demonstrate the empirical power and size of our test, also an application to biological data will be presented.

Test For High Dimensional Data Under Sparsity

✦Ping-Shou Zhong, Iowa State University, 1218 Snedecor Hall,, Departmemt of Statistics, Ames, IA 50011, *pszhong@iastate.edu*; Ping-Shou Zhong, Iowa State University

Key Words: High dimensional data, Large p, small n, Sparsity

In this study, we consider detecting sparse signal in the high dimensional but small sample size data. We propose a threshold test statistic for simultaneous tests on the high dimensional mean under the sparsity condition. The asymptotic distributions of the test statistic are derived for the non-normal and dependent data under the "large p, small n" scenarios. We show that the proposed test could achieve the optimal detection boundary under weak dependence assumption. It was also found that the proposed test performs better than the high dimensional test proposed by Chen and Qin (2010) under the sparsity condition.

Analysis Of Iterative Hard Thresholding For Support Recovery In The High-Dimensional And Noisy Case For Gaussian X Matrices

✦Antony Joseph, Yale University, RM 120, 420 Temple ST, New Haven, CT 06511 US, *antony.joseph@yale.edu*; Andrew Barron, Yale University

Key Words: high dimensional regression, greedy algorithms, Lasso, AWGN channel, compressed sensing

Abstract: Problems of high dimensional regression are becoming increasingly popular nowadays. Greedy algorithms such as Forward Stepwise Regression and Orthogonal Greedy Algorithm (OGA) are commonly used. These algorithms are known to not necessarily select the best subset of explanatory variables. Nevertheless, recent results have shown that they can perform well in function approximation and prediction. It is only very recently that Greedy algorithm began to be analyzed for support recovery with control on pair-wise correlations between variables even when the number of variables p is large compared to the sample size n. Here we analyze a variant of the OGA. We get sharp controls on the relationships between sample size, dimension and sparsity for exact or near exact support recovery. The results specialize to a high dimensional regression setup for a communications system to perform at rates arbitrarily close to capacity for a Gaussian noise channel.

Test For High Dimensional Covariance Matrices With Banded Structure

✦Yumou Qiu, Iowa State University, 4211 Lincoln Swing st Unit 20, Ames, IA 50014, *qiuyumou@gmail.com*; Song Xi Chen, Iowa State University

Key Words: High data dimension, Banded covariance structure, Large p, small n

The paper proposes a nonparametric test for the banded structure of high dimensional covariance matrices, which corresponds to the situations where the variables far apart in the ordering are uncorrelated. We do not assume any specific parametric distribution for the data. The test is robust under the situations where the data dimension is much larger than the sample size, namely the "large p, small n" situations. Specifically, the method can be used to test whether the covariance matrices are diagonal, which means all the variables are mutually uncorrelated. Also, the paper proposes a method to estimate the bandwidth of a covariance matrix which is truly banded. We demonstrate by both theoretical and simulation studies that the test has good properties for a wide range of dimensions and sample sizes.

Halfline Tests For Multivariate One-Sided Alternatives

✦Zeng-Hua Lu, University of South Australia, , *zen.lu@unisa.edu.au*

Key Words: Inequality constraint, One-sided hypothesis, positive orthant, somewhere most powerful test

The idea of using a single linear combination of individual coordinates of a multivariate random variable as the test statistic for testing one-sided hypotheses was suggested in the early literature. This paper extends such a halfline test for testing positive orthant alternatives in the cases of normal distributions with known and unknown covariance, and non-normal distributions with unknown covariance. We first establish a necessary and sufficient condition for the existence of the most stringent somewhere most powerful test proposed in the early literature. When the condition fails, we propose a procedure for implementing halfline tests. We show that our proposed tests possess consistency and (asymptotically) unbiasedness properties, which may have been misleading in the literature. Our simulation studies reveal that the proposed tests can perform better than some existing tests in some situations and are more robust against the normality assumption than the likelihood ratio-based approach. A further appeal of our test is that, unlike many tests for one-sided testing problems, the proposed tests are very easy to implement and generally applicable.

479 Factor Models and Instrumental Variables

Business and Economic Statistics Section

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Sparse Variable Pca With A Vector L_O Penalty

✦Victor Solo, University of New South Wales, Sydney, International 2073 Australia, *v.solo@unsw.edu.au*

Key Words: sparse PCA, l_o, penalised likelihood

By penalising a PCA based likelihood with a novel vector l_o penalty we develop a procedure called sparse variable PCA that first removes a large number of variables and then does a PCA of the remaining variables. We illustrate the procedure with a neuroimaging example although we offer comments on the application to dynamic factor modeling.

Dynamic Common Factors In The Presence Of Seasonality

✦Fabio Humberto Nieto, Universidad Nacional de Colombia, Carrera 30 No. 45-03, Bogot·, 2414 COLOMBIA, *fhnietos@unal.edu.co*; Daniel PeÒa, Universidad Carlos III de Madrid; Dagoberto Saboy·, Universidad Externado de Colombia

Key Words: Dynamic common factors, Multivariate time series, Seasonality

Dynamic common factors, as stochastic processes, have been analyzed under the assumptions that some of them are nonstationary and the others are stationary. Methodologically, this assumption implies that observed time series must be deseasonalized before exploring for common factors in the multivariate time series they conform. Besides the frequent use of determining common factors for getting efficiency in forecasting, they might be used for computing coincident indexes in economy and, from them, leading indexes. It is well known that a deseasonalized time series may exhibit spurious characteristics, specifically, in showing cycles the original data do not contain. In this paper, we address the problem of determining common factors when some of the time series under study are seasonal and this characteristic can be specified by seasonal common factors, in the sense they are seasonal stochastic processes. In this way, one would not need to deseasonalize time series before doing an entertained analysis.

Euromind: Main Features, Achievements And Future Challenges

✦Gian Luigi Mazzi, EUROSTAT, Bech Building, 5, Rue Alphonse Weicker, Luxembourg, International L-2721 Luxembourg, *Gianluigi.Mazzi@ec.europa.eu*; Cecilia Frale, Italian Ministry of the Economy and Finance; Massimiliano Marcellino, European University Institute; Tommaso Proietti, Universit· di Roma Tor Vergata

Key Words: temporal disaggregation, dynamic factor model, coincident indicator, state space methods, national accounts, chain-linking

EuroMIND is a monthly indicator of economic activity for the Euro Area. It provides an estimate of GDP based on temporal disaggregation of quarterly National Accounts estimates, within a set of linked medium-size dynamic factor models for a set of coincident indicators. Main features are the following: first it considers both the output and the expenditure sides of GDP, combining the two estimates with optimal weights; second, monthly and quarterly information is modelled in a unique state-space setup; third, it pays attention to computational efficiency, implementing univariate filtering and smoothing procedures; finally, chain linking and its implications is taken into account. From its original development EuroMIND has been extended over several directions: EuroMIND-S, which includes business survey indicators in a separate factor from other indicators for increasing the now-casting properties; EuroMIND-C, which extends the estimates at member state level; EuroMIND-B proposing estimates over the historical period staring from January 1971; and EuroMIND-G which aims at providing an estimate of potential output and output gap for the Euro area.

Factor Selection In Hedge Fund Replication Dynamic Models: An Application Of Forward Filtering-Backward Sampling Algorithm And Reversible-Jump Mcmc

✦Guillaume Weisang, Bentley University, 9 Melrose St, Apt 2, Boston, MA 02116, *gweisang@bentley.edu*

Key Words: dynamic models, reversible-jump MCMC, Kalman Filter, factor selection, forward-filtering backward-sampling, hedge fund replication

While the use of dynamic factor models for Hedge Fund Replication (HFR) has proven to be superior to standard OLS methodologies (e.g., Roncalli and Weisang, 2009a,b), current factor selection methodology in these dynamic settings by means of direct PCA-based estimation of the factors (e.g., Darolles and Mero, 2008) provides results having poor interpretability and small, yet significant, improvement over more naive selection methodologies. In this paper, I contend on the one hand that this type of direct factor estimation is doomed to be at best biased and unreliable in finite samples as indicated by results from random matrix theory (Harding, 2008). On the other hand, forward filtering-backward sampling (FFBS) algorithm and reversible-jump MCMC allow considering a factor selection methodology using observable instruments as potential factors hence solving both problems of identification and interpretability. I provide here a detailed algorithm and the results of its application to the replication of major hedge fund indices.

Measuring The Uncertainty Of Misspecified Unobserved Components

Andrew Harvey, University of Cambridge; ✦Alejandro Federico Rodriguez, Universidad de ConcepciÛn, Avda. Esteban Iturra s/n - Barrio Universitario, ConcepciÛn, 4030000 Chile, *aleferodriguez@udec.cl*; Esther Ruiz, Universidad Carlos III de Madrid

Key Words: Bootstrap, Conditional heterocedasticity, Iintegrated random walk, Kalman filter, State space models, Stationary and wild bootstrap

In the context of time series analysis, the Kalman filter (KF) is a very powerful tool to estimate underlying components as common factors, trends, time-varying parameters etc. However, the KF requires knowledge of the true specification and parameters. This is unrealistic situation as both the specification and the parameters are often unknown. Consequently, the Prediction Mean Squared Errors (PMSE) associated with the estimated unobserved components is biased. This paper has two contributions. First, we measure the biases attributed to model misspecification in two different contexts. First, we consider the biases when measuring the uncertainty associated with the underlying trend estimated by running the Integrated Random Walk when the true model is a Random Walk plus drift model. Secondly, we consider the biases incurred when the underlying level of the Local Level Model is estimated by assuming homoscedasticity when the true model is conditionally heteroscedastic. The second contribution of this paper is to show that traditional bootstrap procedures designed to obtain the PMSE in the presence of parameter uncertainty are not able to cope with model misspecification.

Information Criteria For Selecting Instrumental Variables In Conditional Moment Restriction Models

✦Gunce Eryuruk, ITAM, Ave. Camino a Santa Teresa #930, Mexico City, 10700 Mexico, *gunce.eryuruk@itam.mx*; Bruce E Hansen, University of Wisconsin

Key Words: generalized method of moments, mean squared error, information criterion, instrument selection, moment selection

The efficiency of GMM estimation is sensitive to the choice of instruments. Increasing the number of instruments decreases the asymptotic variance but increases the finite sample bias of the GMM estimator. A major breakthrough is the work of Donald, Imbens and Newey (2009), who have derived a general expression for the higher-order asymptotic mean square error (MSE) of GMM estimators, allowing for nonlinear simultaneous models with unknown heteroskedasticity. They also suggested an estimator of the MSE to be used as a finite-sample instrument selection criterion, but did not fully justify or investigate this criterion. Our focus is on the construction of an appropriate finite-sample instrument selection criterion. First, we generalize their MSE expression to allow general weight matrices. Second, using asymptotic expansions we formally investigate the bias in the natural plug-in estimator of the MSE. Third, we propose and evaluate bias-corrections for the MSE estimators, in analogy to the classic Mallows criterion for regression models. We investigate the performance of our proposed instrument selection criterion via asymptotic expansions and finite sample simulations.

The Consistent Method Of Moments Estimator In Weak Instrument Cases

✦Ignacio Lobato, Instituto TecnolÛgico AutÛnomo de MÈxico, 10700 MÈxico, *ilobato@itam.mx*; Eduardo GarcÌa GÛmez-Tagle, ITAM

Key Words: consistency, weak instruments, method of moments, instrumental variables

This article studies the behavior of the consistent method of moments (CMM) estimator introduced in Dominguez and Lobato (2004) in the context of weak instruments for both linear and nonlinear cases. For the linear case when many instruments are available, we propose a two step estimator based on a modification of the CMM estimator that is simple to calculate and presents better finite sample performance. For the case where the parameters of interest are defined by nonlinear conditional moments, the CMM estimator is more robust in situations such as weak instruments, where some orthogonality conditions are close to zero. The reason behind this robustness is that the CMM estimator relies on an infinite number of orthogonal conditions to identify these parameters, in contrast to the traditional instrumental variable estimator. We illustrate this robustness with simulations in both linear and nonlinear cases, which compare the finite sample behavior of instrumental variable estimators and our proposed extensions of the CMM estimator.

480 Environmental Monitoring ■●

Section on Statistics and the Environment, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

An Investigation Of Us Temperature Patterns Since 1960?

✦ronald Tracy, oakland University, Department of Economics, Rochester Hills, MI 48309, *tracy@oakland.edu*

Key Words: Temperature, Global Warming, Binomial Test

Global warming has caused the average temperature in the US to rise one degree Fahrenheit since 1970 according to the US National Climatic Center. Because this increase is small relative to temperature variability and because the average temperature is actually the average of the daily low and high temperature (a midrange), standard statistical tests for the mean cannot be used to confirm a systematic temperature increase in the US. Using National Data Climatic Center Version 2 data for the US, this study demonstrates using a binomial test that the US has experienced 50 years of statistically significant temperature increases (Type I error was less than 0.01). Furthermore this study shows that the decade of the 1960s was characterized by statistically significantly cooler weather than the base period and that each decade since then has been characterized by statistically significant warmer weather.

Spatio-Temporal Patterns And Variation Of Common Pm Metals

✦Boubakari Ibrahimou, Western Kentucky university, 1293 Sun Ct, Apt G, Bowling Green, KY 42104, *bibrahim@mail.usf.edu*; Yiliang Zhu, University of South Florida

Key Words: Particulate matter, Metal speciation, Spatio-temporal, Modeling

Exposure to fine particulate matter (PM) in the ambient air is associated with various health effects. There is increasing evidence that the chemical components such as heavy metals of PM play a central role leading to adverse health effects. Thus, there have been intensified efforts in studying the metal composition and it emission origin in ambient PM as well as its association with health outcomes. In this paper we use non parametric smoothing methods to study the spatio-temporal patterns and variation of common PM metals, and illustrate our modeling using data from multiple stations of the EPA's National Air Quality Monitoring Network in three metropolitan statistical areas (MSA). Our goal is to use the spatial variation to estimate chronic and acute residential exposure to these ambient metals, and also use the temporal variation to identify emission sources that are responsible for high ambient concentrations of certain metals while controlling for temperature and barometric pressure. Metals such as aluminum, calcium, iron, chromium, lead, magnesium and zinc show a clear trend or seasonal variations across the three MSAs. Overall we see a decreasing trend for most metals in

Multipollutant Accountability Research Using Principal Stratification

✦Corwin Zigler, Harvard School of Public Health, MA 02115, *czigler@hsph.harvard.edu*; Francesca Dominici, Harvard School of Public Health

Key Words: Accountability research, Causal inference, Principal stratification, Spatial statistics, Regulation policy

Quantifying health effects of actions taken to improve air quality is challenging for large-scale regulations meant to improve air quality and health over a long period of time and over large geographical areas. We propose a potential-outcomes framework for accountability research that adopts ideas rooted in principal stratification to stratify areas based on the estimated causal effect of an air-quality regulation on ambient air pollution and assess the causal effect of the regulation on health outcomes within these strata. Because regulations may affect concentrations of many interrelated pollutants, we extend the use of principal stratification to a multipollutant approach that accommodates a continuously-scaled multivariate intermediate response vector. Furthermore, we make use of recent advancements in hierarchical modeling for point-referenced spatial data to capitalize on information contained in the geographic locations of the pollution measurement sites. Insofar as these methods rely on unobserved potential outcomes, they rely on assumptions regarding associations not identified based on observed data, indicating the importance of sensitivity analyses.

Optimized Endogenous Post-Stratification In Forest Inventories

✦Paul L Patterson, US Forest Service, 2150A Centre Ave, Suite 350, Fort Collins, CO 80525, *PLPatterson@fs.fed.us*

Key Words: endogenous post-stratification, environmental monitoring, forest inventory, remote sensing, post-stratification, logistic regression

An example of endogenous post -stratification is the use of remote sensing data and a sample of ground data to build a logistic regression model to predict the probability that a plot is forested and then use the predicted probabilities to form categories used for post-stratification. Breidt and Opsomer showed design consistency of the endogenous post-stratified estimator and, under a super-population model, the consistency and asymptotic normality of the endogenous post-stratified estimator by showing it has the same asymptotic variance as the traditional post-stratified estimator with prespecified strata. An optimized endogenous post-stratified estimator where the optimization occurs over all cut points has been recently proposed in the literature. There are no known literature results describing the operating characteristics of this new estimator. This study reports the results of a detailed Monte Carlo investigation of the small and large sample performance of the optimized endogenous post-stratified estimator under a variety of realistic scenarios and compares its performance with earlier approaches. The results provide guidance as to the appropriateness of these estimators.

Communicating Environmental Health Risk In Uncertain Times

✦Margaret G. Conomos, United States Environmental Protection Agency, 1200 Pennsylvania Avenue, N.W., 2842T, Washington, DC 20460 USA, *conomos.margaret@epa.gov*; Barry D. Nussbaum, United States Environmental Protection Agency

Key Words: risk communication, risk assessment, role of statistician, health hazards, toxic releases

The public has a very understandable desire to assess their health and safety in light of environmental hazards. The Environmental Protection Agency has a compendium of tools that provide risk assessments. The risks are defined in terms of scores, numbers, reports, or guidance. Do these tools truly answer the public's questions? How well does the public translate this information to evaluate how safe they are? This presentation will describe several approaches to answer questions that the public has about their safety with regard to exposures to chemicals. The paper will emphasize that the role of the statistician does not end with analysis and evaluation; but rather the statistician has a crucial role in the clear communication of risk.

Trends In Minimum Temperatures

✦Lynne Seymour, The University of Georgia, Department of Statistics, 101 Cedar St., Athens, GA 30602-7952, *seymour@uga.edu*; Carolanne Lott, The University of Georgia; Julia Orr, The University of Georgia; David Stooksbury, The University of Georgia

Key Words: climate change, time series

We investigate trends for three different statistics on daily minimum temperatures: the monthly mean, the monthly standard deviation, and the monthly maximum of these recorded values. The data are periodically correlated, and trends across the state of Georgia will be examined.

Modeling Aspen Tree Mortality Using Modified Gee Approach For Irregular Time Intervals

✦Suborna Shekhor Ahmed, University of British Columbia, 2424 Main Mall, Forest Sciences Centre, Forest Resources Management Department, UBC, Vancouver, BC V6T1Z4 Canada, *suborna.ahmed@gmail.com*; Valerie LeMay, University of British Columbia; Changhui Peng, University of Quebec at Montreal; Zhihai Ma, University of Quebec at Montreal

Key Words: irregular time intervals, modified GEE, maximum likelihood estimator, modified logistic regression, survival analysis, autocorrelation

A probability of survival model was used to predict regular mortality of the aspen (Populus tremuloides Michx) tree species in Alberta, Canada. For tree survival, commonly permanent sample plot (PSP) data with repeated measures of individual trees at irregular time intervals are used. As a result, the logistic model was modified for survival under irregular time intervals (Modified Logistic Model). Also, autocorrelation within trees under irregular time intervals was addressed by developing a modified Generalized Estimating Equations (GEE) approach. In the Modified GEE approach, the error correlation structure was adjusted using the tree-specific time intervals. The Maximum Likelihood estimator for the Modified Logistic Model assuming no autocorrelation was used to obtain the initial estimate of the vector of model parameters. Then, the Modified Logistic Model was fit using the modified GEE approach to obtain final estimates. Monte Carlo simulations indicate that the resulting estimators are approximately unbiased and are reasonably precise.

481 Bayesian Statistical Methodology with Epidemiologic Applications

Section on Statistics in Epidemiology, Section on Bayesian Statistical Science

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

On Non-Parametric Bayesian Regression In Cardiovascular Disease Risk Assessment

✦Olli Saarela, McGill University, Department of Epidemiology, Biostatistics and OH, Purvis Hall, 1020 Pine Avenue West, Montreal, QC H3A 1A2 Canada, *olli.saarela@thl.fi*; Elja Arjas, University of Helsinki

Key Words: disease prediction, risk assessment, non-parametric Bayesian regression, continuous model expansion, monotonic regression, cardiovascular diseases

Assessing the absolute risk for a future disease event for presently healthy individuals has an important role in primary prevention of cardiovascular diseases (CVD). In addition to utilizing novel risk factor information on top of well established risk factors, the risk assessments can potentially be improved by using better, more realistic, statistical models. Here we concentrate on this latter task, utilizing non-parametric Bayesian regression techniques and continuous model expansion to relax parametric modeling assumptions where this is most beneficial. While allowing for non-linear associations all around, to achieve parsimony in the model fit, we allow multidimensional relationships within specified subsets of risk factors, determined either on a priori basis or as part of the estimation procedure. The computational tool utilized is our previously published Bayesian multivariate monotonic regression procedure, which we generalize here to a survival analysis setting. We apply the proposed methods to study whether 10-year CVD risk assessment in Finnish population based cohorts can be improved by more careful modeling of the effects of the classic risk factors of CVD.

Bayesian Inference Of Odds Ratios In Misclassified Binary Data With A Validation Substudy And An Application To Cervical Cancer Data

✦Dewi Gabriela Rahardja, UT Southwestern Medical Center, TX USA, *rahardja@gmail.com*; Yan Daniel Zhao, UT Southwestern Medical Center; Hao Helen Zhang, North Carolina State University

Key Words: Bayesian inference, Binary data, Credible interval, Case-control study, Misclassification, Odds ratio

We propose a fully Bayesian model with a non-informative prior for analyzing misclassified binary data with a validation substudy. In addition, we derive a closed-form algorithm for sampling all parameters from the posterior distribution and making statistical inference on odds ratios. Our algorithm draws each parameter from a beta distribution, avoids the specification of initial values, and does not have convergence issues. For illustration, we consider a case-control study which explores the association between herpes simplex virus and invasive cervical cancer. We then apply our algorithm to this case-control and compare the results with those obtained by other existing methods. Finally, we assess the performance of our algorithm using simulation studies.

Risk Evaluation For Obesity Related Diseases Using Bayesian Hierarchical Model

✦Casey Jelsema, Western Michigan University, 1018 Claymoor Dr Apt 1C, Kalamazoo, MI 49009 United States, *casey.m.jelsema@wmich.edu*; Rajib Paul, Western Michigan University; Joseph D'Ambrosio, Michigan State University / Kalamazoo Center for Medical Studies; Luis Toledo, Michigan State University / Kalamazoo Center for Medical Studies

Key Words: CAR Model, Center for Disease Control, Cardiovascular Diseases, MCMC

In recent years, obesity related diseases such as diabetes, hypertension, cardiovascular diseases, etc. have become prevalent in the United States. We develop a Bayesian hierarchical model to characterize these multiple disease interactions. Our model incorporates the effects of an individual's lifestyle, pre-existing medical conditions, other demographic information. We fit our model to the collected by the Center for Disease Control as a part their Behavioral Risk Factor Surveillance System. We produce disease-risk maps. Inferences on the model parameters are based on Markov Chain Monte Carlo simulation.

A Bayesian Model For Estimation Of Hiv Incidence Using Serial Prevalence Data

✦Nicole Bohme Carnegie, New York University, 246 Greene St, 3rd Floor, New York, NY 10003, *nicole.carnegie@gmail.com*

Key Words: HIV incidence, serial prevalence data, Bayesian modelling

One of the most common HIV monitoring systems in resource-limited settings is antenatal or neo-natal seroprevalence surveillance. Under this system, we have prevalence estimates in women of child-bearing age repeated, usually annually, over an extended period of time. We can then attempt to extract incidence rates over time, based on changes in population prevalence and known factors influencing the likelihood of an infected person being tested, for example excess mortality or disease-related infertility. We build on the flexible method for estimating incidence from serial prevalence data, using a relative inclusion ratio (RIR) to adjust for differential inclusion of infected and uninfected persons in the sample developed by Ades and Medley (1994). We reformulate the model in a Bayesian framework and extend it to include adjustment for the sensitivity and specificity of tests. The resulting model is applied to data from King County, Washington, with a new specification of the RIR for application to surveillance data in public health clinics.

Estimating Cancer Mortality Rates Using Bayesian Area-Age-Period-Cohort (Aapc) Model And Multistage Carcinogenesis Models

✦Zhiheng Xu, Emory University, 6200 Crestgate Lane, Tucker, GA 30084 USA, *zxu4@emory.edu*; Vicki S. Hertzberg, Emory University

Key Words: AAPC, carcinogenesis, DIC

AAPC models have been widely used as a general framework to study the temporal and spatial pattern of the risk of disease. It is well known that there is a non-identifiability problem associated with the age, period and cohort effects because of the exact linear relationship among them. Bayesian procedures have been implemented to handle the identifiability issues in the APC models through prior setting and model constraints. However, some constraints lack of a strong biological explanation. In this study, we plan to introduce the multistage carcinogenesis models into AAPC model to incorporate more biological meaning to the age effects. The age effect in the AAPC model will be replaced by the hazard function derived from carcinogenesis model and new priors will be assigned on the carcinogenesis model. Spatial correlation will be taken into account in the AAPC model. Colon cancer death rate in IOWA will be used in this study and we will apply spatial effect to the county level. Deviance Information Criteria (DIC) will be used to compare different models in fitting cancer mortality data.

Modeling Of Exposure Data From Point Sources Under A Matched Case-Control Design With Multiple Disease States

✦Shi Li, Department of Biostatistics, University of michigan, 4413 Center Valley Dr, Ann Arbor, MI 48108 USA, *shili@umich.edu*

Key Words: matched case-control study, point source modeling, disease sub-classification, Bayesian inference, Markov chain Monte Carlo (MCMC), conditional likelihood

In this paper, we extend the distance-odds models of Diggle et al (2000) proposed to investigate elevated risk/odds around point sources under a matched case-control design to distance-odds models where there are sub-types within cases in a matched design. We consider models analogous to the polychotomous logit models and adjacent-category logit models. Multiple point sources as well as covariate adjustments are considered. Maximum likelihood, profile likelihood, iteratively re-weighted least squares and a hierarchical Bayesian approach using Markov chain Monte Carlo were evaluated to conduct inference under these models. We compare these methods via an extensive simulation study and show that with multiple parameters and a nonlinear model, Bayesian methods have advantages in terms of estimation stability, precision and inference. The methods are illustrated by analyzing Medicaid claims data in pediatric asthma population in Detroit.

482 Novel Statistical Approaches for Analyzing Gene-array Data

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

Wednesday, August 3, 8:30 a.m.–10:20 a.m.

Power And Stability Comparisons Of Multiple Testing Procedures With Fdr Control

✦Dongmei Li, University of Hawaii, , *dongmeil@hawaii.edu*

Key Words: Multiple testing, FDR, Power, Stability, Adaptive procedure

Microarray analysis has been widely used to identify differentially expressed genes under different treatment conditions or to identify gene markers associated with certain disease or quantitative traits. False discovery rate (FDR) is considered to be a proper type I error rate to control for microarray data analysis. Various multiple testing procedures are proposed to control the FDR, which include several adaptive procedures incorporating the estimated proportion of true null hypotheses in the testing procedures. However, the stability of those multiple testing procedures with FDR control has not been evaluated yet. To determine which procedure most likely to produce reproducible results, simulations were conducted to study the stabilities of five frequently used multiple testing procedures (BH 1995, BY 2001, Storey2002, Storey2004, BKY 2001) with FDR control. Our simulation results showed the adaptive linear step-up procedure and modified version, proposed by Storey (2002, 2004), performs best considering both power and stability. An ovarian cancer microarray study is used to illustrate the stability of those five multiple testing procedures with FDR control.

An Alternative Method Of Pair-Wise Sequence Alignment

Mian Arif Shams Adnan, JAHANGIRNAGAR UNIVERSITY; ✦Md. Moinuddin, JAHANGIRNAGAR UNIVERSITY, M. S. Student, DEPARTMENT OF STATISTICS, SAVAR, DHAKA, International 1342 BANGLADESH, *akmmusju@yahoo.com*; SHONGKOUR ROY, JAHANGIRNAGAR UNIVERSITY; Md. Riashad Jaman, JAHANGIRNAGAR UNIVERSITY

Key Words: Gap penalty, Sequence Alignment, Transition probability

All previously derived methods of pair-wise sequence alignments suffer from either deciding the appropriate gap penalty or not ensuring more accuracy. This paper inaugurated a new method for pair-wise sequence alignment which overcomes both of the aforementioned problems. This method is based on the differences between transition probabilities.

Reducing Type I Error Rate In Association Study Of Whole-Genome Sequencing Data

✦Renfang Jiang, Michigan Technological University, Department of Mathematical Sciences, Houghton, MI 49931, *rjiang@mtu.edu*; Jianping Dong, Michigan Technological University

Key Words: wavelet transform, quantitative traits, multiple testing, moving windows, rare variants, collapsing method

We conducted a genome-wide association study (GWAS) on a simulated unrelated data using a multilocus score test based on wavelet transformation proposed recently by the authors. Wavelet transformation is an advanced smoothing technique while the currently popular collapsing methods are the simplest way of smoothing multilocus genotype. The wavelet-based test suppressed noise from the data more effectively, which results in lower type I error rates. A level dependent threshold is chosen for the wavelet-based test to suppress the optimal amount of noise according to the data. Multiple remedies are proposed in this paper to reduce inflated type I error rate including: to use a window of fixed size rather than a gene; to use Bonferroni correction rather than comparing to maximums of test values for multiple testing correction; to remove the influence of other factors by using residuals for association test. A wavelet-based test can detect multiple rare functional variants. Type I error rates can be controlled using wavelet-based test combined with the remedies mentioned above.

Snpmclust: A Model-Based Genotype-Clustering And -Calling Algorithm For Illumina Genotyping Microarrays

✦Stephen Erickson, University of Arkansas for Medical Sciences, Arkansas Center for Birth Defects Research, 13 Children's Way #512-40, Little Rock, AR 72202 USA, *serickson@uams.edu*

Key Words: microarrays, genotyping, genomics

Genotyping microarrays ("SNP chips") produce measurements of fluorescent signal intensity corresponding to the alleles of each probed single nucleotide polymorphism (SNP). We introduce SNPMClust, a bivariate Gaussian model-based genotype-clustering and -calling algorithm for Illumina SNP chips, which builds on MClust, a contributed package in the R statistical computing environment, and GenCall, Illumina's proprietary genotyping algorithm. SNPMClust has been used in multiple studies at the Arkansas Center for Birth Defects Research & Prevention. The accuracy and sensitivity of SNPMClust were compared with those of GenCall on a dataset of 96 cheek swab DNA samples run on a custom 1536-SNP Illumina GoldenGate array, at 1156 SNPs for which the true genotype was known with high confidence. The area under the ROC curve is 4.3% greater for SNPMClust than GenCall, and the ROC curve for SNPMClust is uniformly higher than that for GenCall at false call rates lower than 2%. SNPMClust, therefore, provides a useful research tool, either as the primary source of genotype calls or as a diagnostic tool.

Background Correction Based On The Box-Cox Transformation Of Noises For Illumina Bead Array Data

✦Min Chen, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-9066 United States of America, *Min.Chen@UTSouthwestern.edu*; Yang Xie, University of Texas Southwestern Medical Center

Key Words: background adjustment, Box-Cox Transformation, Illumina Bead Array

Abstract The Illumina BeadArray becomes increasingly popular due to many attractive features. One distinction with other platforms is that it has negative control beads containing arbitrary oligonucleotide sequences that are not specific to any target genes in the genome. This design provides a way of directly estimating the distribution of the background noise. In the literature of background correction, the noise is often assumed to be normal. However, we show with real data that the noise can be very skewed, and the correction methods based on the normality assumption can lead to biased gene expression intensities. In this study we propose a noise adjustment method based on a model with a Box-Cox transformation on the noise term. To facilitate the search of MLE estimators, a spline technique is applied to approximate the likelihood function, and we show it can greatly improve the performance of the searching algorithm.

Modeling Long Time-Series Gene Expression Data With The Kolmogorov-Zurbenko Algorithm

✦Darlene M Olsen, Norwich University, 158 Harmon Drive, Northfield, VT 05663, *dolsen1@norwich.edu*

Key Words: micorarray, time series, spline

Long time-series microarray analysis offers an exciting opportunity in biomedical research to investigate the progression of temporal gene expression profiles after certain intervention or treatment, which can yield a more accurate assessment of how altered cellular pathways may interact over time and give insight to the progression of diseases. The Kolmogorov-Zurbenko algorithm uses the Kolmogorov-Zurbenko Spline (KZS) as the modeling framework for temporal profiles and clusters genes based on the parameters of the KZS function. Applying this statistical methodology to analyze time-series microarray data will improve our cognizance of biological networks.

Ultrasensitive Detection Of Rare Variants In Admixture Populations Using Next-Generation Sequencing

✦Patrick Flaherty, Stanford University, 318 Campus Dr. W330, Stanford, CA 94027, *pflahert@stanford.edu*; Georges Natsoulis, Stanford University; Omkar Muralidharan, Stanford University; Nancy Zhang, Stanford University; Jason Buenrostro, Stanford University; John Bell, Stanford University; Mark Winters, Stanford University; Mark Holodniy, VA Palo Alto Health Care System; Hanlee Ji, Stanford University

Key Words: next-generation sequencing, genomics, viral, hierarchical model

The viral load of an infected individual often contains a population of quasi-species each with a different genome. A rare resistant variant may survive lethal selection and proliferate causing prolonged illness. We describe improvements to both statistical and experimental methodology using next-generation sequencing to identify rare mutational events. We develop a hierarchical model to account for over-dispersion in the observed data. We show a theoretical power analysis involving detection limit improvements due to both sample preparation precision and read depth. We present results on the detection of a quasi-species present at a prevalence of 0.1% with high sensitivity (100%) and specificity (99%) using an experimentally defined admixed population of synthesized DNA.

483 Likelihood, nuisance parameter and estimation efficiency ●

International Chinese Statistical Association, International Indian Statistical Association, SSC

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

A Dimension-Reduction Approach for Generalized Linear Models

✦Lixing Zhu, Hong Kong Baptist University, Kowloon Tong, Hong Kong, NA P. R. China, *lzhu@hkbu.edu.hk*

Key Words: quasi-likelihood, Least squares, Fisher information, sparse model, predictor selection

Quasi-likelihood is one of the most popularly used methods for estimating parameters in generalized linear models with asymptotic efficiency. However, when the dimension of predictor vector is large, solutions of the corresponding estimating equations are unstable and even not convergent. This is clearly the case for sparse models in ``large p, small n" paradigms. In this paper, we propose a two-stage estimation approach. Different from classical quasi-likelihood, we first employ linear least squares for transformed response to obtain an estimation for a vector proportional to the parameter of interest. We then use quasi-likelihood to estimate one-dimensional scale of the parameter. As linear least squares is of a very simple closed form and is very efficient in computation, we can then efficiently reduce original dimension of predictor vector to one first so that we can efficiently apply quasi-likelihood to estimate scale of parameter. When the transformation for response is bounded, the new estimation is robust against distribution of error. The method is applied to predictor selection.

Jackknife Empirical Likelihood For Case-Control Study with Auxiliary Information

✦Bing-Yi Jing, Hong Kong University of Science and Technology, Clear Water Bay,, Kowloon, Hong Kong, , *majing@ust.hk*

Key Words: Jackknife, empirical likelihood, case-control

In this talk, we propose a jackknife empirical likelihood method to do inference for the interested parameters of the multiplicative-intercept risk models taking into account auxiliary information on controls in case-control studies. It is shown that the proposed statistic asymptotically follows a chi-squared distribution in this set-up. Simulation studies investigate the small-sample properties. This is joint work with Zhouping LI and Wang ZHOU.

Let's Practice What We Preach: Likelihood Methods for Monte Carlo Data

✦XIAOLI MENG, Harvard University, Department of Statistics, 1 Oxford Street, 7th Floor, Cambridge, MA 02138 USA, *meng@stat.harvard.edu*

Key Words: Bridge Sampling, Empirical Likelihood, MCMC, Monte Carlo Integration, Normalizing Constant, Semi-parametric models

A central challenge in statistical inferences with real data is that all models are wrong. Ironically, when we apply likelihood methods to Monte Carlo data, we face the opposite problem of knowing too much because they are simulated from a completely known model. Simply treating our direct computational ``estimand", such as a Bayes factor, as our ``unknown" model parameter turns out to be inadequate. This talk outlines recent progresses on a semi-parametric likelihood approach proposed by Kong, McCullagh, Meng, Nicolae, and Tan (2003, JRSSB), which formulated this inference problem by treating the underlying baseline measure as the model parameter and by modeling how much known information about the measure is ignored by the Monte Carlo simulation. All Monte Carlo integrals are then estimated as linear functionals of the MLE of the baseline measure. This approach not only shows that standard methods such as importance sampling and more generally bridge sampling are MLEs and thus they are the best possible ones for the amount information we typically ignore, but more importantly it leads to new estimators, if we allow ourselves to ignore less, that are more efficient.

484 Innovative Applications of Non- or Semiparametric Methods in the Fields of Biology and Medicines ■●

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

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Inverse Regression for Functional and Longitudinal Data

✦Jane-Ling Wang, University of California, Davis, Department of Statistics, Davis, CA 95616, *wang@wald.ucdavis.edu*; Ciren Jiang, Statistics and Applied Mathematical Sciences Institute; John Aston, University of Warwick

Key Words: dimension reduction, functional data analysis, principal component analysis, smoothing

Sliced inverse regression is an appealing dimension reduction method that was developed for regression models with multivariate covariates. The aim of this talk is to extend it to both functional and intermittently measured longitudinal covariates. Previous work in dimension reduction with functional covariates concentrates on the case where the whole trajectories of random functional covariates are assumed to be observed completely. Our approach allows for longitudinal covariates that are recorded discretely and intermittently. We develop asymptotic theory for the new procedure and show that the effective dimension reduction space can be estimated at the root-n rate of convergence. The effectiveness of the approach is demonstrated through a fecundity data of Medflies.

Clustering and Variable Selection for Functional Data

✦Gerda Claeskens, Katholieke Universiteit Leuven, OR & Business Statistics, Naamsestraat 69, Leuven, 3000 Belgium, *gerda.claeskens@econ.kuleuven.be*

Key Words: Functional data, clustering, variable selection, warping

A first part of the talk concerns a quantification of phase and amplitude variability in a sample of curves. As opposed to treating phase variation as a nuisance effect, we explicitly recognize it as a possible important source of information for clustering, or classification. Another part of the talk is devoted to variable selection. We consider functional regression models with both vector and functional covariates. It is of interest to select the relevant variables to simplify the model. This requires semiparametric variable selection methods.

Corrected-Loss Estimation for Quantile Regression with Covariate Measurement Error

✦Huixia Judy Wang, North Carolina State University, 2311 Stinson Drive, 4270 SAS Hall, Raleigh, NC 27695, *judy_wang@ncsu.edu*; LEONARD A. STEFANSKI, North Carolina State University; Zhongyi Zhu, Fudan University

Key Words: Corrected loss function, Laplace, Measurement error, Normal, Quantile regression, Smoothing

We study estimation in quantile regression when covariates are measured with error. Existing work in the literature often requires stringent assumptions, such as spherically symmetric joint distribution of the regression and measurement error variables, or linearity of all quantile functions, which restrict model flexibility and complicates computation. In this paper, we develop a new estimation approach based on corrected scores to account for a class of covariate measurement errors in quantile regression. The proposed method is simple to implement, and its validity only requires linearity of the particular quantile function of interest. In addition, the proposed method does not require any parametric assumptions on the regression error distributions. We demonstrate with simulation study that the proposed estimators are more efficient than existing methods in various models considered. Finally we illustrate the proposed method through the analysis of a dietary data.

Efficient Association Tests for Gene/SNP Sets/Networks

✦Xihong Lin, Harvard School of Public Health, 655 Huntington Ave, Harvard School of Public Health, Building 2, Rm 419, Boston, MA 02115 USA, *xlin@hsph.harvard.edu*; Tianxi Cai, Harvard University; Raymond James Carroll, Texas A&M University

We develop an efficient test for studying the association of disease phenotypes and gene/SNP sets or gene networks. This class of tests accounts for the fact that genes within a gene set/network are often correlated and the signals are likely to be sparse. The proposed test effectively dampens the effects of the noisy variables. We study the theoretical properties of the test and show that they are more powerful than the traditional multivariate tests especially when the gene set is large. We apply the proposed methods to genome-wide association studies to study the effects of genomic features on disease phenotypes.

485 Recent Method Development on High-dimensional Statistical Learning

Section on Statistical Learning and Data Mining, International Indian Statistical Association, Section on Statistical Computing

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Necessary and Sufficient Conditions Toward Optimal Feature Selection and Parameter Estimation

✦Xiaotong Shen, The University of Minnesota, School of Statistics, 313 Ford Hall, Minneapolis, MN 55455 USA, *shenx002@umn.edu*; Wei Pan, University of Minnesota; Yunzhang Zhu, University of Minnesota

Key Words: High-dimensional regression, Nonconvex regularization

High-dimensional feature selection has become increasingly crucial for seeking parsimonious models in parameter estimation. For selection consistency, we derive one necessary and sufficient condition formulated on the notion of degree-of-separation. The minimal degree of separation is required for consistency, and consistency is achieved by constrained L0-regularization and its computational surrogate, at the degree of separation slightly exceeding the minimal level. This permits up to exponentially many features in the sample size, regardless of if a true model exists. In contrast, their unconstrained counterparts do so in presence of a true model, as in the parametric case. In this sense, L0-regularization and its surrogate are optimal in feature selection against any method. More importantly, sharper parameter estimation and prediction are resulted from such selection, which, otherwise, is impossible in absence of a good selection rule.

A Road to Classification in High-Dimensional Space

Jianqing Fan, Princeton University; ✦Yang Feng, Columbia University, Room 1012, 1255 Amsterdam Ave, 10th Floor, MC 4690, New York, NY 10027, *yangfeng@stat.columbia.edu*; Xin Tong, Princeton University

Key Words: ROAD, RAID, high dimensional classification, Fisher Rule, Independence Rule

For high-dimensional classification, researchers proposed independence rules to circumvent the diverse spectra, and sparse independence rule to mitigate the issue of noise accumulation. However, in biological applications, there are often a group of correlated genes responsible for clinical outcomes, and the use of the covariance information can significantly reduce misclassification rates. The extent of such error rate reductions is unveiled by comparing the misclassification rates of the Fisher discriminant rule and the independence rule. To materialize the gain based on finite samples, a Regularized Optimal Affine Discriminant (ROAD) is proposed based on a covariance penalty. Inspired by ROAD, a Regularized Affine Independence Discriminant (RAID) is also proposed to improve independence rules. An efficient constrained coordinate-wise descent algorithm (CCD) is also developed to solve the optimization problem associated with the ROAD and RAID. Oracle type of sampling properties are established. Simulation studies and real data analysis support our theoretical results and demonstrate the advantages of the new classification procedure under a variety of correlation structures.

A Second Look at Stability Selection

✦Richard Samworth, University of Cambridge, Statistical Laboratory, Wilberforce Road, Cambridge, CB3 0WB United Kingdom, *r.samworth@statslab.cam.ac.uk*; Rajen Shah, University of Cambridge

Key Words: Variable selection, subsamples, bagging

We study the Stability Selection procedure for variable selection introduced in the recent JRSSB discussion paper by Meinshausen and Buhlmann (2010). Stability Selection is essentially `Bagging with error control', and we show that a minor variant of the original procedure yields control of the false selection rate even for finitely many subsamples. In fact, a much improved bound holds under reasonable additional conditions. This allows the practitioner to reduce the threshold for the proportion of subsamples in which a variable must be selected in order to be declared significant, which is helpful because the original procedure tends to select rather few variables. We also discuss other aspects and applications of this promising methodology.

The Sparse Laplacian Shrinkage Estimator for High-Dimensional Regression

✦Jian Huang, University of Iowa, Department of Statistics and Actuarial Science, 241 SH, Iowa City, IA 52242, *jian-huang@uiowa.edu*; Shuangge Ma, Yale University; Hongzhe Li, University of Pennsylvania; Cun-Hui Zhang, Rutgers University

Key Words: Graphical structure, Minimax concave penalty, Penalized regression, High-dimensional data, Variable selection, Oracle property

We propose a new penalized method for variable selection and estimation that explicitly incorporates the correlation patterns among predictors. This method is based on a combination of the minimax concave penalty and Laplacian quadratic associated with a graph as the penalty function. We call it the sparse Laplacian shrinkage (SLS) method. The SLS uses the minimax concave penalty for encouraging sparsity and Laplacian quadratic penalty for promoting smoothness among coefficients associated with the correlated predictors. The SLS has a generalized grouping property with respect to the graph represented by the Laplacian quadratic. In a special case, it has a similar grouping property as the elastic net method. We show that the SLS possesses an oracle property in the sense that it is selection consistent and equal to the oracle Laplacian shrinkage estimator with high probability. This result holds in sparse, high-dimensional settings with $p \gg n$ under reasonable conditions. We derive a coordinate descent algorithm for computing the SLS estimates. Simulation studies are conducted to evaluate the performance of the SLS method and a data example is used to illustrate its application

486 Using advanced visual analytics to improve safety assessment and decision making in drug development ■

Section on Health Policy Statistics, ENAR, Section on Risk Analysis, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Safety Graphics: Graphical Approaches for Identifying Safety Signals

✦Janelle Charles, Office of Biostatistics,OTS, CDER, FDA, Division of Biometrics VII, 10903 New Hampshire Avenue - Building 21 Room 4619, Silver Spring, 20993, *janelle.charles@fda.hhs.gov*

Key Words: safety, graphics

Currently most statistical summaries of safety information from clinical trials are displayed in lengthy tables or listings. It is challenging for both reviewers and decision makers to use these presentations to quickly identify safety signals, to adequately characterize the safety profile of pharmaceutical products in the context of their intended indication, and ultimately to make decisions regarding benefit-risk. The purpose of this presentation is to investigate a set of graphics that aid safety experts in identifying and evaluating safety information more effectively and efficiently.

Creating Effective Visual Tools for the Assessment and Characterization of Pharmaceutical Product Safety: General Principles, Illustrations, and Public Access

✦Kenneth J. Koury, Merck Research Laboratories, 2015 Galloping Hill Road, K-15-2,2050, Kenilworth, NJ 07033, *kenneth.koury@merck.com*

Key Words: visualization tools, FDA/Industry Safety Graphics Working Group, CTSpedia

General principles for creating effective visualization tools will be reviewed, and examples which illustrate the effective and efficient communication of safety information based on clinical trial data obtained during the product development program will presented and discussed. The use of the CTSpedia website will also be described. This site allows public access to all of the visualization tools developed by the FDA/Industry Safety Graphics Working Group, and it provides an opportunity for users to contribute new graphics or suggest modifications for existing graphics.

Visualization Tools and the Evaluation of Safety Data in Pharmaceutical Industry Clinical Trials

✦Qi Jiang, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, *qjiang@amgen.com*

Key Words: Safety Data, Clinical Trials, Safety Signals, Visualization Tools

Recent events have put a spotlight on the evaluation of safety data from pharmaceutical industry clinical trials. The evaluation of safety information from clinical trials has always been challenging for a number of reasons. These include the frequent lack of prespecified hypotheses, making it extremely difficult to distinguish true safety signals from spurious ones, and the fact that some events are quite rare and can only be detected in large trials of long duration. In this presentation we will discuss some proposals for addressing these problems through improved visualization tools.

Visually Displaying Benefit:Risk:Uncertainty in Safety Assessments

✦Mark O Walderhaug, U.S. Food and Drug Administration, 20852, *Mark.Walderhaug@fda.hhs.gov*; Richard A Forshee, U.S. Food and Drug Administration, CBER; Arianna Simonetti, U.S. Food and Drug Administration, CBER; Anne Fernando, Norfolk State University

Key Words: benefit, risk, safety, efficacy, visual display, uncertainty

The visual display of safety information is critical to all stakeholders in the process of medical decision-making. While safety information is vital, visualizing benefits and the uncertainty of benefits along with safety information can help keep the decision-making balanced. The visual display of the results of Phase III studies can show a relative equality in the uncertainty of both benefits and risks; however, post-market data streams are asymmetric and can show a more rapid decrease in risk uncertainty because while safety information is reported, post-market efficacy data is generally not reported. As the uncertainty about risk is reduced, the magnitude of risk is likely to increase as low frequency adverse events are reported over time. The ultimate goal of drug/biologic safety decision-making is greater public health, and the visual display of both benefits and risks along with uncertainties helps keep this public health goal in the decision-maker's mind.

487 Session in Honor of Professor Gary Koch on his Festschrift ■●

Statistics in Biopharmaceutical Research Journal, Biopharmaceutical Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Investigating Identity Matching in GWAS: Potholes in the Path to Personalized Medicine

◆J. Richard Landis, University of Pennsylvania School of Med., Dept. of Biostatistics & Epidemiology, 423 Guardian Drive, 600 Blockley Hall, Philadelphia, PA 19104-6021, *jrlandis@mail.med.upenn.edu*

Key Words: forensic biostatistics, category-specific intraclass correlation, identity misalignment, categorical data, agreement statistics

This talk will illustrate the use of forensic biostatistics methods to investigate identity mis-alignment of genotype to phenotype data. Within a multicenter clinical research network, n=3,668 participants with genetic testing consent were selected for genotyping. Alerted by PID mismatches of duplicate SNP data, category-specific intraclass correlation coefficients (Landis et al 2010) comparing self-reported and genetically-inferred race were implemented within subsets of PIDs known to have been processed separately, and the presence of non-random clustering of race disagreement patterns was discovered on selected genotyping plates. Well established associations between HDL-C and CETP gene SNPs were shown to be seriously attenuated (approx. 20%) within race subgroups. A subsequent GWAS fingerprinting sub-study, focusing on 24 identity SNPs common to both studies, permitted re-alignment of these identities, so that discovery and validation research could proceed. Because of the confounding of race in HDL-C and SNP associations, these results illustrate the major impact that identity mis-alignment can have on biomarker discovery and/or validation.

Genomic Classifier for Patient Enrichment: Misclassification and Type I Error Issues in Pharmacogenomics Noninferiority Trial

◆Sue-Jane Wang, Office of Biostatistics, CDER, US FDA, WO#21, Mailstop 3562, 10903 New Hampshire Ave., Silver Spring, MD 20993, *Suejane.Wang@fda.hhs.gov*

Key Words: Biomarker Qualification, Disease prevalence, False positive classification, Nondifferential misclassification, noninferiority, positive predictive value

Unlike the phenotypic characteristics, a genomic diagnostic assay is needed to assess whether a patient truly has the genomic characteristics of interest for therapeutic investigation. However, in reality, the diagnostic assay is not perfect without misclassification error. We consider a targeted or enrichment pharmacogenomics clinical trial with a non-inferiority objective. In this paper presentation, we assess the maximum type I error rate associated with falsely concluding non-inferiority in the presence of misclassification of subject's genomic status. To achieve a prescribed level of positive predictive value, the feasibility of a genomic biomarker for enrichment in a non-inferiority study is a function of disease prevalence. If the evidential standard of much less inflation of type I error rate can be achieved, a genomic biomarker, if qualified, may be useful as an adjunct tool for patient selection.

Inferiority Index and Margin Specification in One-Sided Equivalence Studies

◆George YH Chi, J&J Pharmaceutical Research and Development, L.L.C., 920 US Highway 202 S, Raritan, NJ 08869, *gchi@its.jnj.com*; Gang Li, J&J Pharmaceutical Research and Development, L.L.C.

Key Words: inferiority index, margin, degree of tightness, non-inferiority trial, comparative study

In (one-sided) equivalence studies, the setting of an equivalence margin has often been controversial. This is true for example in non-inferiority trials. Much of the problems can be traced to a lack of an objective standard for measuring the stringency of a margin. In this presentation, we will introduce the concept of an inferiority index. It will be shown that under different types of distributions, the inferiority index can be linked to a given effect measure of interest. Through this link, the inferiority index can then provide a measure of the degree of tightness of a given margin for the effect measure. In addition, through this link, a margin can be specified with the desired level of stringency, although in practice, this should be done in conjunction with other relevant information. Its applications to non-inferiority trials will be illustrated with examples from survival, normal and Bernoulli distributions. Potential applications to other comparative studies will be noted.

488 Meeting Challenges for Modeling Brain Imaging Data: The Spatio-Temporal Perspective ■●

ENAR, International Indian Statistical Association, Committee of Representatives to AAAS

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Utilizing Induced Voxel Correlation in fMRI Analysis

◆Daniel B. Rowe, Marquette University, 1313 W. Wisconsin Ave., Milwaukee, WI 53233, *daniel.rowe@marquette.edu*

Key Words: fMRI, spatio-temporal, imaging, preprocessing

In functional magnetic resonance imaging (fMRI), it is known that there exists spatio-temporal correlation between voxels over time. Attempts have been made to model voxel spatio-temporal correlations in fMRI data. In fMRI, preprocessing of images and voxel time series are almost always performed to "improve" the data. It has been shown that these preprocessing steps can modify the mean image and induce

spatio-temporal correlation between voxels that would be uncorrelated without preprocessing. Methods are described to utilize the known correlation from preprocessing that is induced for statistical analysis.

Meeting Challenges for Modeling Brain Imaging Data: The Spatio-Temporal Perspective

✦DuBois Bowman, Emory University, 1518 Clifton Rd., N.E., 3rd Floor, Rollins School of Public Health, Emory Univ., Atlanta, GA 30322, *dbowma3@emory.edu*; Shuo Chen, Emory University

Key Words: Imaging, spatial modeling, fMRI, resting-state analysis, mental health

Analyzing resting-state fMRI data has revealed brain networks that exhibit consistent properties across subjects. Moreover, altered resting-state network properties are associated with mental illnesses, including major depressive disorder (MDD). Most resting-state fMRI studies conduct network analyses; however, important information may exist in localized "activation" properties. A key challenge in characterizing localized activation properties for resting-state fMRI data is that first-order moments are not interpretable as activation statistics. We extend a Bayesian spatial model (Bowman et al., 2008) that was primarily intended for the analysis of task-induced neural processing, which yields inferences for both localized activation patterns and network properties. We first construct frequency-domain descriptors of the resting-state activity profiles, then use a Bayesian hierarchical model to estimate localized patterns of this measure as well as a covariance matrix reflecting the functional connectivity between brain regions. We perform estimation using MCMC implemented via Gibbs sampler. We apply our Bayesian model to data from a resting-state fMRI study of MDD.

A Bayesian Spatiotemporal Model for Multi-Subject EEG

✦Wesley Thompson, University of California, San Diego, 9500 Gilman Drive 0664, La Jolla, CA 92093-0664 USA, *wes.stat@gmail.com*

Key Words: EEG, Multi-Subject, Bayesian, Functional Neuroimaging, Source-Localized

Electroencephalography (EEG) has millisecond temporal resolution but localization of neural sources inside the brain is a difficult problem. One approach is to apply independent components analysis to scalp EEG recordings to reduce the dimensionality of sources, and then to use a forward head model to determine the best-fitting source dipole location. Applied to different subjects, this typically results in varying numbers and locations of source dipoles across subjects. We develop a Bayesian spatiotemporal model for multi-subject source-localized EEG which exploits time-frequency domain relationships to obtain group inferences on the spatial locations and dynamic inter-relationships among localized sources. Each subject's localized sources is modeled as arising from a mixture distribution of spatial coordinates and time-varying partial coherence. Model inference is obtained via a Markov Chain Monte Carlo algorithm. The utility of this method is demonstrated by simulations and by application to a large dataset of event-related EEG.

489 Advances in measurement error adjustment in nutrition and physical activity epidemiology ■●

Section on Statistics in Epidemiology, ENAR, WNAR

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Measurement Error Correction in Nutritional Epidemiology for Survival Data Analysis with Time-Varying Exposures

Donna Spiegelman, Harvard School of Public Health; ✦Xiaomei Liao, Harvard School of Public Health, , *stxia@channing.harvard.edu*; David Zucker, Hebrew University; Yi Li, Dana-Farber Cancer Institute and Harvard School of Public Health

Key Words: survival analysis, measurement error, nutrition, epidemiology

Nutritional epidemiologists are often interested in estimating the prospective effect of time-varying exposure variables such as cumulative exposure or cumulative updated average exposure, in relation to chronic disease endpoints such as cancer incidence and mortality. By re-calibrating the measurement error model within each risk set, a risk set regression calibration (RRC) method is proposed for this setting. An algorithm for a bias-corrected point estimate of the relative risk using an RRC approach is presented, followed by the derivation of an estimate of its variance. Emphasis is on methods applicable to the main study/external validation study design, which is standard in nutritional epidemiology, and where the exposure of interest is a cumulative average. Simulation studies under several assumptions about the error model were conducted, which demonstrated the validity and efficiency of the method in finite samples. The method was applied to a study of long-term vitamin E and calcium intake in relation to colorectal cancer risk and to a study of long-term diet quality in relation to diabetes risk, both in Harvard's Health Professionals Follow-up Study.

Modeling the Measurement Error Structure of Physical Activity Data

✦Sarah Nusser, Iowa State University, Dept of Statistics, Ames, IA 50011-1210, *nusser@iastate.edu*

Key Words: measurement error, physical activity

Activity data take many forms (e.g., energy expenditure, metabolic equivalents (METS), time spent sedentary activities) and are often collected in epidemiological studies. Although most physical activity data collection methods generate data with substantive measurement error, statistical error adjustments are not yet commonly applied to activity data when making inferences. As with dietary intake, the concept of "usual" activity is of interest when making inferences about long-run average behaviors for individuals and about the relationship of activity to chronic health outcomes. Using data from a survey in which respondents provided duplicate observations of concurrent activity recall and activity monitor data, we will present models to estimate error characteristics of recall and sensor data measurements. We will also use a measurement error framework to estimate the distribution of usual activity for subpopulations to characterize behaviors at the population level.

Advances in Measurement Error Adjustment in Nutrition and Physical Activity Epidemiology

✦Ross L Prentice, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M3 B873, Seattle, WA 98109, *rprentic@whi.org*

Key Words: measurement error, nutrition, physical activity, regression calibration, hazard ratio, systematic bias

On the Use of Biomarker Sub-Studies to Correct for Measurement Error in Nutrition and Physical Activity Epidemiology Ross l. Prentice Fred Hutchinson Cancer Research Center and University of Washington Biomarker studies of nutrient consumption and activity-related energy expenditure will be described.These include a Nutrient Biomarker Study among 544 postmenopausal women in the Women's Health Initiative Dietary Modification Trial, and 450 women in the WHI prospective cohort study.A doubly-labeled water(DLW)biomarker of energy consumption, a urinary nitrogen biomarker of protein consumption and a (0.9*DLW -resting energy expenditure) objective assessment of activity-related energy expenditure will be log-transformed and regressed on corresponding log(self-report consumption/activity estimates) and such study subject characteristics as body mass, age and ethnicity, to yield calibrated consumption/activity estimates. Disease association estimates will be shown to be quire different and stronge following biomarker calibration, for cancer and cardiovascular diseases. Ongoing work to develop additional nutitional biomarkers via a human feeding study will be briefly described.

Modeling Longitudinal Observations with Excess Zeros and Measurement Error, with Application to Nutritional Epidemiology

✦Victor Kipnis, National Cancer Institute, Biometry Research Group, 6130 Executive Boulevard, EPN 3131, MSC 7354, Bethesda, MD 20892, *kipnisv@mail.nih.gov*; Raymond James Carroll, Texas A&M University; Laurence S Freedman, Gertner Institute for Epidemiology and Health Policy Research; Douglas Midthune, Biometry, National Cancer Institute

Key Words: regression calibration, episodically consumed foods, correction for measurement error

Food frequency questionnaires have been the instrument of choice to assess long-term dietary intake in most cohort studies of diet and disease. It is now well appreciated that such questionnaires involve substantial measurement error, random and systematic, leading to distorted effects of diet. A popular method for correcting for this error, regression calibration, assumes a substudy with an unbiased short-term reference instrument. Application of this method to foods that are not consumed every day by everyone is problematic, since short-term reference instruments usually include a substantial proportion of subjects with zero intakes, leading to observations with excess zeros, in addition to within-person measurement error. Also, dietary intake is often analyzed relative to total energy intake to take into account dietary composition. We present a new bivariate model for short-term reference observations on food and total energy intakes with application to calibration of energy-adjusted associations between foods and disease. We exemplify the methodology by applying it to data from the US NIH-AARP Diet and Health Study and study its properties using simulations.

490 Active surveillance in observational data: lessons from the OMOP project ■

IMS

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

A Generalization of the Self-Controlled Case Series Method Allowing for Positive Event Dependence

✦Shawn E. Simpson, Columbia University, New York, NY 10027, *shawn@stat.columbia.edu*

Key Words: recurrent events, self-controlled case series, drug safety, event dependence, conditional Poisson model

The self-controlled case series (SCCS) method is used to analyze recurrent events and determine their association with time-varying covariates. It is based on a conditional Poisson regression model, which assumes that events at different time points are conditionally independent given the covariate process. This is particularly problematic when the occurrence of an event can alter the future event risk. In a clinical setting, for example, patients who have a first myocardial infarction may be at higher subsequent risk for a second. In this work we propose a generalization of SCCS that allows the occurrence of an event to increase the future event risk, yet maintains the advantages of the original model by controlling for fixed baseline covariates and relying solely on data from cases. We will focus on the application of post-marketing drug safety surveillance and examine how well our method performs in discerning the relationship between drug exposures and adverse health outcomes.

Performance of a Comparative New User Design for Drug Safety Signal Detection

✦Maurice Alan Brookhart, UNC Chapel Hill, Department of Epidemiology, McGavran-Greenberg, CB # 7435, Chapel Hill, NC 27514 USA, *abrookhart@unc.edu*

We describe a general approach for detecting drug safety signals in large administrative databases based on a comparative new user design combined with propensity score (PS) methods to adjust for confounding by baseline covariates. We propose and empirically compare various drug safety signal detection assays based on different implementations of this general approach. The implementations vary in how the cohort is defined (cohort entry occurring at first use versus a run-in period); how the PS is specified (i.e., what covariates are included); whether the PS is truncated to eliminate non-overlap; and how follow-up and censoring are handled. We then discuss the performance of this approach across a range of known or strongly suspected drug-outcome relations or non-relations.

Operating Characteristics of Statistical Methods for Active Drug Safety Surveillance

✦Patrick Ryan, Johnson & Johnson, , *ryan@omop.org*

The Observational Medical Outcomes Partnership (OMOP, http://omop.fnih.org) is a public-private partnership conducting methodological research to inform the appropriate use of stastical methods when applied to observational healthcare databases for active drug safety surveillance. OMOP designed an experiment to study the performance of different active surveillance methods against an array of disparate observational databases, in their ability to identify known drug associations and discern from false positive findings. Thirteen methods were implemented as alternative analysis approaches for active surveillance, including cohort and case-based designs, sequential testing approaches and multivariate models used in other data mining applications. These methods were executed across the OMOP data network of 10 data sources, both administrative claims and electronic health records, and run for a set of drug-outcome test cases with a defined ground truth. The resulting estimates were used to measure the operating characteristics (e.g. sensitivity, specificity) of the methods across the individual sources and within composite estimates obtained from meta-analysis.

491 Modern Treatments of Semiparametric Inference ●

Section on Nonparametric Statistics, International Chinese Statistical Association

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Semiparametric Estimation in a Single Index Model with Endogenous Variables

✦Ingrid Van Keilegom, UniversitÈ catholique de Louvain, Institute of Statistics, Voie du Roman Pays 20, Louvain-la-Neuve, 1348 Belgium, *ingrid.vankeilegom@uclouvain.be*; Melanie Birke, Ruhr-Universit‰t Bochum; SÈbastien Van Bellegem, UniversitÈ catholique de Louvain

Key Words: endogeneity, single-index model, inverse problem, instrumental variable, efficiency

In this talk we consider the estimation of a semiparametric single-index model when endogeneity is present. The presence of an instrument is assumed that is non-correlated with the error term. The proposed estimator is the solution of an inverse problem. The asymptotic properties of the proposed estimator are studied, and the estimators are also studied for small samples through simulations.

Semiparametric Optimality Under Nonstandard Conditions

✦Michael R Kosorok, UNC-CH, Department of Biostatistics,, UNC-CH, Chapel Hill, NC 27599, *kosorok@unc.edu*

Key Words: semiparametric optimality, non-identifiability, constrained estimation

We consider semiparametric optimality under two non-standard conditions. The first condition happens when the semiparametric model has a parameter that is not identifiable under a null hypothesis being tested. The second condition happens when there are a finite number of smooth inequality constraints on the parameters. In both settings, careful attention must be paid to local asymptotic normality. This is particularly true in the first setting where the appropriate limiting distribution is a Gaussian process rather than the usual Gaussian vector. In the second setting, minimaxity and admissibility for the constrained estimator boil down to minimaxity and admissibility for a single, multivariate normal observation with mean constrained to a cone and variance unknown. Unfortunately, finding either minimax or admissible estimators for this single normal observation case is currently an unsolved problem. Recent developments on these problems will be discussed along with several open questions.

Robust Regression Using Unimodal Density Estimation

✦Mary C Meyer, Colorado State University, 212 Statistics Building, Fort Collins, CO 80523 USA, *meyer@stat.colostate.edu*

Key Words: contaminated data, influential points, heavy-tailed errors, outliers, nonparametric density estimation

The error density for a regression model is assumed to be unimodal and symmetric with mean zero. The error density and regression parameters are estimated simultaneously, allowing for thin- or heavy-tailed errors. Inference about the parameters uses the estimated density.

Nonparametric Regression with Continuous and Categorical Predictors

✦Jean Opsomer, Colorado State University, CO 80523, *jopsomer@stat.colostate.edu*; Jay Breidt, Colorado State University; Ismael Sanchez, University of Granada

Key Words: kernel regression, small area estimation

We will consider the problem of fitting a nonparametric regression model for the case of a continuous dependent variable and a mixture of continuous, ordinal and nominal independent variables. We will use generalized product kernel regression, which combines kernels appropriate for all three types of covariates. Some asymptotic theory for the ordinal variable kernel regression will be developed and applied to this scenario. This work is motivated by a small area estimation problem of the U.S. Bureau of Labor Statistics, who are interested in estimating the mean wage by job type, job level and geographic area based on data from the National Compensation Survey.

492 Surveys, Bayesian Modeling, and the Interconnections: A Session in Honor of Joseph Sedransk ■●

Section on Survey Research Methods, Section on Government Statistics, Section on Bayesian Statistical Science, International Indian Statistical Association

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Estimation of Finite Population Domain Means: Then and Now

✦Jiming Jiang, University of California, Davis, Department of Statistics, University of California, Davis, CA 95616 USA, *jiang@wald.ucdavis.edu*; Thuan Nguyen, Oregon Health and Science University

Key Words: Asymptotics, Mean Square Prediction Error, Model Misspecification, Nested-Error Regression, Robustness, Small Area Estimation

The first part of the talk is based on a paper published in JASA(2006; joint with P. Lahiri), in which we introduce a general methodology for producing a model-assisted empirical best predictor (EBP) of a finite population domain mean using data from a complex survey. Our method improves on the commonly used design-consistent survey estimator by using a suitable mixed model. Unlike a purely model-based EBP, the proposed model-assisted EBP converges in probability to the customary design-consistent estimator as the domain and sample sizes increase. The convergence is shown to hold with respect to the sampling design, irrespective of the assumed mixed model, a property commonly known as design-consistency. The second part of the talk introduces a new approach to small area estimation. We derive the best predictive estimator (BPE) of the fixed parameters under the nested-error regression model. This leads to a new prediction procedure, called observed best prediction (OBP). We show that the OBP is more reasonable than the empirical best linear unbiased predictor (EBLUP) and can significantly outperform the latter when the underlying model is misspecified.

Calibrated Bayes Inference for Sample Surveys

✦Rod Little, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109 USA, *rlittle@umich.edu*

Key Words: Calibrated Bayes, survey sampling, design based inference

The calibrated Bayesian approach to surveys basis inferences on the posterior predictive distribution from a Bayesian model, but seeks inferences that are frequency calibrated, in the sense of have good frequentist properties in repeated sampling. The calibrated aspect dictates that survey design features like clustering and weighting need to be appropriately included in the model, to protect against model misspecification. Examples are provided on how to do this, and calibrated Bayes inferences are shown by simulation to have better frequentist properties than standard design based approaches, including the model-assisted approach popular with current survey practitioners.

Bayesian Predictive Inference for Small Areas with Uncertain Unit-Level Models

✦Donald Malec, National Center for Health Statistics, 3311 Toledo Rd , Room 3210, Hyattsvillle, 20782, *DMalec@cdc.gov*; Ryan Janicki, U. S. Census Bureau

Validation of a unit level model at the small area level may be difficult due to small sample size. In addition, the sample design may often be informative at this level and needs to be accounted for. Lastly, Bayesian inference requires a complete likelihood, necessitating the need for a detailed description of the small area likelihood. The consequences of knowing little about the sampling distribution beyond the first two moments, including the dependence between observations will be demonstrated.Criteria for choosing distributions will be discussed.

Bayesian Predictive Inference for a Proportion with Nonresponse

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

Key Words: Gibbs sampler, Log-linear, Multinomial-Dirichlet, Patterns of missingness, Sensitivity analysis

In any survey an individual may respond to one or more items, and generally it is not possible to determine whether a nonrespondent differs from a respondent with the same traits. In the Slovenian Public Opinion Survey (SPOS) with the three categorical variables (secession, attendance, independence), there are eight patterns of missingness (eight tables), and the emphasis was in determining the proportion of the population who would attend the plebiscite and vote for independence. One strategy is to center a nonignorable nonresponse model on an ignorable nonresponse model, thereby adding uncertainty to the ignorable nonresponse model. The categorical tables are model simultaneously with different parameters, and these parameters are independent and identically distributed. They differ by one or more sensitivity (or centering) parameters, whose prior specifications give the difference between the ignorable and nonignorable nonresponse models. With these specifications, a Bayesian approach is attractive, and the computation can be done using the Gibbs sampler. Thus, it is possible to study imprecision and ignorance to obtain an uncertainty interval. The SPOS data provide an illus

493 Statistics in Computational Advertising ■●

Section on Statistical Computing, Section on Statistical Graphics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Time-Sensitive Collaborative Filtering via Item-Item Similarities

✦Rahul Mazumder, Stanford University, Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305, *rahulm@stanford.edu*; Deepak Agarwal, Yahoo!; Liang Zhang, Yahoo!

Key Words: collaborative filtering, covariance regularization, sparsity (L1 regularization), large scale optimization, expectation maximization, web / internet applications

Consider a generic collaborative system problem with users rating (clicking) items over time --- with potentially a large fraction of the user-item matrix unobserved. We will address how to come up with predictions/ recommendations for the unobserved entries in the matrix by incorporating item-item similarity in a model-based fashion. This novel proposal takes us well beyond the typical low-rank matrix factorization models used in collaborative filtering. Various notions of item-item similarities can be used via a prior regularization eg the sparse inverse covariance regularization accounting for interpretable conditional independencies among the items resulting in a sparse item-item similarity graph and sophisticated variants thereof. We propose computationally tractable, theoretically justified, novel algorithms that enable us to learn these complex models for large-scale recommender systems, with potentially billions of variables. We will also discuss how adopt our framework to account for dynamic user /item visit patterns and time sensitive user-item interactions, as arsing in Yahoo! front page and News applications.

Predicting Response Rates in Ad Exchanges

✦Deepak Agarwal, Yahoo!, 4401 Great America parkway, santa clara, CA 95054 USA, *dagarwal@yahoo-inc.com*

Key Words: computation advertising, spike and slab prior, Gamma-Poisson, multi-level hierarchical model, map-reduce, ad exchange

An ad exchange provides a unified marketplace to connect buyers (advertisers) and sellers (publishers,) possibly through commercial intermediaries like ad-networks or publisher networks. One key statistical challenge for proper functioning of such an exchange involve estimating response rates of rare events like click on an ad, user buying a product on publisher page and so on. I will present a statistical model that estimates such response rates through a multi-level hierarchical model. Our model performs shrinkage by leveraging data aggregates along multiple hierarchies and scales to massive data sets in a map-reduce framework (billions of observations, hundreds of millions of covariates).

Issues in Targeted Advertising for Local Search

✦Suhrid Balakrishnan, AT&T Labs Research, 180 Park Ave., Florham Park, NJ 07932 USA, *suhrid@research.att.com*

Key Words: computational advertising, CTR modeling, local search

Advances have recently been made on modeling click-through rate in well studied settings like sponsored search and context match. Local search, has received relatively less attention. The geographic nature of local search and associated local browsing makes interesting research challenges and opportunities possible. We consider a novel application of a relational regression model to local search. The model is attractive in that it allows us to explicitly control and represent geographic and category-based neighborhood style constraints on the samples that result in superior click-through rate estimates.

User Browsing Models: Relevance Versus Examination

✦Sugato Basu, Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043 USA, *sugato@google.com*

Key Words: browsing models, relevance, examination

There has been considerable work on user browsing models for search engine results, both organic and sponsored. The click-through rate of a result is the product of the probability of examination (will the user look at the result) times the perceived relevance of the result (probability of a click given examination). Past papers have assumed that when the CTR of a result varies based on the pattern of clicks in prior positions, this variation is solely due to changes in the probability of examination. We show that, for sponsored search results, a substantial portion of the change in CTR when conditioned on prior clicks is due to a change in the relevance of results for that query instance, not just due to a change in the probability of examination. We then propose three new user browsing models, which attribute CTR changes solely to changes in relevance, solely to changes in examination, or to both changes in relevance and examination. The model that attributes all the CTR change to relevance yields substantially better predictors of CTR than models that attribute all the change to examination, and does only slightly worse in predicting relevance.

494 Medallion Lecture: Loss, Actions, Decisions: Bayesian Analysis in High-Throughput Genomics ■●

IMS, ENAR, International Indian Statistical Association

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Loss, Actions, and Decisions: Subjective Inference in High-Throughput Genomics

✦Chris Holmes, University of Oxford, Department of Statistics, Oxford, OX1 3TG U.K., *cholmes@stats.ox.ac.uk*

Key Words: genomics, Bayesian inference, decision theory, nonparametrics

In applications of high-dimensional data analyses arising in genetics and genomics we often have substantive prior knowledge on plausible effect sizes and on the credible dependence structures underlying the data. Subjective Bayesian analysis that exploits this knowledge and incorporates losses on action-outcome pairs is particularly well suited here. In this talk we will overview recent developments in Bayesian methods for high-dimensional cluster and regression analysis with an emphasis on nonparametric and decision theoretic approaches in genomics. We will discuss current research in decision theoretic Bayesian inference for partial and nonparametric likelihoods. The methods are motivated and illustrated using ongoing real world studies in cancer genomics, in particular surrounding statistical methods to model the occurrence of copy-number alterations in cancer genomes, whereby stretches of DNA are systematically duplicated or deleted during tumour formation, and their association with clinical phenotypes.

495 Noether Invited Session ●

Noether Award Committee, Section on Nonparametric Statistics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Community Detection and Network Models

✦Liza Levina, University of Michigan, , *elevina@umich.edu*

Analysis of networks and in particular discovering communities in networks has been a focus of recent work in several fields, with applications ranging from friendship networks to food webs. Nonetheless, there is a certain amount of disconnect between the many algorithms proposed for community detection, the probability models for random graphs that are often over-simplified but easily tractable, and the complex features we observe in real networks. The talk will review some common methods and models for community detection, and present new models and methods that allow for more flexible network structures. Both empirical results and asymptotic theory will be used to assess performance.

496 Addressing the Evidence for Anthropogenic Climate Change

ASA Advisory Committee on Climate Change Policy, Scientific and Public Affairs Advisory Committee, Section on Risk Analysis

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Addressing the Evidence for Anthropogenic Climate Change

✦Richard L Smith, University of North Carolina/SAMSI, Department of STOR, Chapel Hill, NC 29599-3260, *rls@email.unc.edu*; ✦L Mark Berliner, Ohio State University, Dept of Statistics, Columbus, OH 43210, *mb@stat.osu.edu*; ✦Ross McKitrick, Department of Economics, University of Guelph, , *rmckitri@uoguelph.ca*; ✦John R. Christy, University of Alabama in Huntsville, Earth System Science Center, NSSTC, 320 Sparkman Drive, NSSTC 4040, Huntsville, 35805, *john.christy@nsstc.uah.edu*

Key Words: Climate Change

The American Statistical Association Board of Directors in 2007 endorsed a statement “Warming of the climate system is unequivocal ... and [m]ost of the observed increase in globally averaged temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations.” In 2009, the ASA president joined leaders of 17 other scientific organizations in signing a letter to U.S. Senators saying, “...climate change is occurring, and rigorous scientific research demonstrates that the greenhouse gases emitted by human activities are the primary driver.” This session, which will included representatives from the ASA Advisory Committee on Climate Change Policy (ACCCP), will examine the evidence for anthropogenic climate change from a statistical viewpoint. The speakers will also discuss the role of statisticians in the climate change science discussions.

497 Teaching Statistics in a Medical School Environment: Different Challenges with Different Audiences ■●

Section on Teaching of Statistics in the Health Sciences, Section on Health Policy Statistics, WNAR

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Teaching Introductory Statistics And Study Design To Residents In A Teaching Hospital

✦Lori Lyn Price, Tufts CTSI, 800 Washington St, Box 63, Boston, MA 02111, *lprice1@tuftsmedicalcenter.org*; Jessica Paulus, Tufts CTSI; John Griffith, Tufts CTSI

Key Words: CTSA, biostatistics, epidemiology, teaching medical residents

The Research Design Center at the Tufts Clinical and Translational Science Institute (CTSI) has led more than 20 series of clinical research seminars across a variety of clinical departments. Each series consists of 5-7 non-credit seminars selected by the department and includes clinical applications of methodologies. Each seminar covers a specific study design or statistical topic and lasts an hour. The teaching objective is to provide clinician participants, mostly residents, with enough information about basic study design and statistics to increase their research comfort level. Online evaluations are collected using 5-point Likert-scale questions (5=best). The mean responses for whether learning objectives were met, material was presented at appropriate technical level, and overall satisfaction are 4.38, 4.22, and 4.33. The short time frame in which to cover material is challenging. Engaging the full set of students is difficult as many residents are required to attend regardless of their interest in the content. Future directions for growth include online seminars and advanced classes taught over several weeks, and increased tracking of the impact of the seminars.

Teaching Statistics in the Dental School Environment: Challenges and Strategies

✦Deborah V Dawson, University of Iowa College of Dentistry, N439E Dental Science Building, U of Iowa College of Dentistry / Dows Institute, Iowa City, IA 52242-1010, *deborah-dawson@uiowa.edu*

Key Words: statistical education, dental education, graduate education, dentistry

Statistical education in the dental school environment must reach varied constituencies, including predoctoral students, residents, graduate students, and faculty. This challenge is complicated by the pressures of a crowded curriculum and the multiple demands made of practitioners in academic dentistry. A critical overarching goal is producing informed consumers of the dental literature, particularly in terms of quantitative literacy. It is also vital that individuals from all of these groups learn how to interact profitably with statisticians in the context of research planning and proposal development, through to the analysis and presentation stages. In addition to formal didactic coursework in statistics and research design, strategies in attaining these goals include integration of statisticians into the predoctoral curriculum in critical thinking and both predoctoral and graduate research programs. Participation in junior faculty development programs and collegiate research seminars, including both general presentations and invited workshops, are used to address faculty needs. Similarly varied approaches are taken to evaluation, including audience response technology.

Teaching Introductory Statistics To Students In The Basic Sciences

✦Constantine Daskalakis, Thomas Jefferson Uninversity, Division of Biostatistics, 1015 Chestnut St, Suite M100, Philadelphia, PA 19107, *c_daskalakis@mail.jci.tju.edu*; Benjamin Leiby, Thomas Jefferson University

Key Words: statistical education, graduate education, introductory statistics, teaching

A two-course sequence of introductory statistics (total of 48 contact hours) was taught to 36 graduate students in basic science programs, during the 2009 and 2010 spring terms. In place of a required text, notes and handouts were written specifically for the class. Stata was used as the software package. Most students had a recipe-type inclination towards statistics, particularly focused on significance. Three quarters of the students pointed to Stata as the most unsatisfactory element of course, expressing strong preference for the use of packages currently used in their labs. A strong minority of the students also ex-

pressed desire for more paper and pencil hand calculations, at least for the beginning of the course. Some students reported that the material taught was too advanced and that they only needed to know enough to be consumers of statistics, but others felt that the course did not cover sufficiently advanced situations relevant to their research work. Students felt that the course goals and objectives were moderately clear (mean = 3.1 in 2009, 3.7 in 2010) and were only partially met (mean = 3.2 and 3.3). Overall satisfaction was low (mean = 2.4 and 2.5).

Teaching An Introductory Biostatistics Sequence In A College Of Nursing At A University Medical Center As A "Satellite" Biostatistician

✦Mary J Kwasny, Northwestern University, 680 N. Lake Shore Dr., #1400, Chicago, IL 60611, *m-kwasny@northwestern.edu*

Teaching a Biostatistics sequence at an academic medical center that does not have a Biostatistics Department, program, or concentration can have unique challenges. For eight academic years between 1999 and 2007, over 25 introductory or mid-level biostatistics courses were taught to more than 700 nursing students in such a setting. With limited biostatistical resources, a single course was designed to accommodate undergraduate, masters, and doctoral level nursing students, as well as students from other disciplines, as need dictated. The nursing program curriculum underwent major changes including decreasing the sequence from three courses to two, and transitioning from in-person to both in-person and on-line formats. Finding and training teaching assistants proved difficult; and common concerns such as the selection of appropriate books and statistical software were not completely understood by the non-statistical faculty. This talk will highlight these issues, discuss processes by which they were resolved, the impact on students, and offer reflections on how some of these issues may have been avoided.

Teaching Basic Biostatistics To Public Health Graduate Students

✦John McGready, Johns Hopkins Bloomberg School, 615 N. Wolfe Street, Baltimore, MD 21215, *jmcgread@jhsph.edu*

For the past 11 years, the two term sequence "Statistical Reasoning in Public Health" has been taught concurrently on campus and online at the Johns Hopkins Bloomberg School of Public Health, with a total enrollment of 300. Initially the primary audience for this course was students pursuing a Master of Public Health (MPH) degree, a group primarily composed of MDs. However, over time the diversity of health related backgrounds in the MPH class has increased. Additionally the audience for the course has grown to encompass students seeking M.S. degrees in specialized areas, including mental health, policy, and social/ behavioral sciences. Originally, the examples used in the course were clinically motivated. However, temporal monitoring of course evaluations illustrated students' desire for data examples that span the breadth of public health. In the most recent overhaul of the course, much effort was put into diversifying the materials to include contemporary examples from the clinical, sociological, epidemiological, policy literature. This talk focuses on sources for diverse examples, and the challenge of addressing heterogeneity of interests in presenting a unified course.

498 Increasing the Impact of Statistical Consulting in Industry using Statistical Engineering ■●

Section on Statistical Consulting, Section on Quality and Productivity, Section for Statistical Programmers and Analysts

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Consulting Through "Statistical Engineering": Moving Up The Value Chain

✦Steven Larson, DuPont, 140 Cypress Station Road, Suite 135, Houston, TX 77090-1633, *steven.f.larson@usa.dupont.com*

Key Words: Consulting, Engineering, Quality, Organizational Impact

The Applied Statistics Group in DuPont resides within its Engineering Department so "statistical engineering" is a natural extension of our function. An early example of statistical engineering as defined in recent literature is that of DuPont's Process Quality Management. As our group has evolved there are many more such examples developed in our consulting practice that have demonstrated the increased value of our group and the application of our discipline on the company as a whole. This talk will discuss the common features of such efforts in consulting which include a strong prevalent business need, collaboration, evolution of the methodology, and, of course, the unique contributions of statisticians.

The Evolution Of Statistical Consulting

✦Robert G Wilkinson, The Lubrizol Corporation, 29400 Lakeland Boulevard, Mail Drop 152B, Wickliffe, OH 44092, *robert.wilkinson@lubrizol.com*

Key Words: statistical engineering, consulting, business process

This is the most exciting time in history to be a statistician. Through advancements in technology, vast quantities of data are being generated and electronically captured on a routine basis. Corporations that master the use of this information will have a significant advantage over those that don't. In addition, technological advancements have resulted in extremely powerful tools for statisticians. These tools enable better analyses. They also provide new ways to communicate the outcomes from the analyses, resulting in a widespread impact on the corporation. Whereas the opportunities are enormous, the challenges are formidable. This talk will focus on the organizational aspects as statistical consulting evolves from providing advice on smaller somewhat isolated problems, to working on the implementation of data-driven systems that have a major impact on key business processes within the corporation. The perspective will be from real-life statistical engineering experiences encountered at Lubrizol.

Statistical Contributions To The Complete Analytical Process

✦Stephanie Palermo, Capital One Financial, 23113 USA, *stephanie.palermo@capitalone.com*

Key Words: engineering, financial, business, innovation

The use of statistics in business has grown tremendously in the past decade. Not only are increasingly more industries and companies employing statisticians, but the application of statistics in those industries is broader and deeper than ever. No longer is statistics used solely for simple analyses or predictions. For many companies, statistical methods are now a part of every step of analysis, including data exploration, test design, forecasting, and monitoring. The process by which discovery occurs as a comprehensive set of advanced analytics and statistics can be described as statistical engineering. Capital One has delivered advances in innovation through the exploration and use of powerful tools and methods. This allows for a more productive partnership with our internal clients and business partners. This talk will focus on the advances made at Capital One and across the industry, and the expanded use modern statistical applications played in those advances.

499 Modern Reliability Data: Opportunities and Challenges ■

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

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Field-Failure Predictions Based On Failure-Time Data With Dynamic Covariate Information

✦William Q. Meeker, Iowa State University, 2109 Snedecor Hall, Department of Statistics, Ames, IA 50011, *wqmeeker@iastate.edu*; Yili Hong, Department of Statistics, Virginia Tech

Key Words: Accelerated failure time model, Lifetime data, Usage history, Cumulative exposure model, Dynamic data, Warranty returns

Today, there are more and more products installed with automatic data-collecting devices such as smart chips and sensors as well as cellular and network communications capabilities that track how and under which environments the product is being used. While there is tremendous amount of such dynamic data being collected, there is little research on using such data to provide more accurate reliability information for products and systems. Motivated by a consulting problem, this paper focuses on using failure-time data with dynamic covariate information to make warranty and other field-failure predictions. The dynamic covariate information is incorporated into a parametric failure-time model through a cumulative exposure model. A prediction procedure that accounts for unit-to-unit and temporal variability in the use rate is developed to predict field-failure returns at a future time. We also define a metric to quantify the improvements obtained using dynamic information in the prediction accuracy. Simulation studies were conducted to study the effect of different sources of covariate process variability on predictions.

Analysis And Modeling Of Data From An Accelerated Destructive Degradation Test With Batch-To-Batch Variability

✦Necip Doganaksoy, General Electric, One Research Circle, K15A60, Niskayuna, NY 12309, *doganaksoy@ge.com*; Ming Li, Applied Statistics Lab, GE Global Research

Key Words: lognormal distribution, reliability, variance components estimation

Accelerated Destructive Degradation Tests (ADDT) are commonly used in industry to speedily gather life and reliability information on new materials and components. Standard statistical models and software can be used for effective analysis of most data sets from ADDTs. In our work involving a new battery component, the initial analysis of the data based on standard ADDT models suggested the presence of an additional source of variability that was not represented in the standard model. In this presentation we will describe our attempt at estimating the contribution of this batch-to-batch variability and its practical implications.

New Perspectives In Analyzing Aircraft Maintenance Data

✦I-Li Lu, The Boeing Company, PO Box 3707 MC 7L-21, Seattle, WA 98124-2207, *i-li.lu@boeing.com*

Key Words: Troubleshooting Procedures, Inventory Management, System Prognostics, Risk Management, Health Management

An Integrated Decision Evaluation & Analysis Library recently developed by Boeing statisticians and maintenance engineers to satisfy statistical analysis requirements for optimization of scheduled maintenance programs will be briefly summarized in this talk. The technology utilizes fleet maintenance and reliability data to determine the most optimum inspection intervals for the scheduled tasks. The statistical analysis methodology can be further expanded to support business needs and opportunities beyond aircraft maintenance industry.

Integrated System Health Management with Reliability and Risk

✦Aparna Huzurbazar, Los Alamos National Laboratory, MS F600, CCS-6 Statistical Sciences Group, Los Alamos, NM 87545, *aparna@lanl.gov*; David H Collins, Los Alamos National Laboratory

Key Words: prognostics, flowgraph model, weapon system, nuclear reactor

Understanding and managing the health of today's complex systems requires a multitude of tools. This talk will give an overview some problems in this area with a focus on Prognostics and Health Management (PHM). The systems that we consider are are typically mission- or safety-critical, expensive to replace, and operate in environments where reliability and cost-effectiveness are a priority. PHM programs are essential for understanding and managing these systems for high reliability at minimum cost. In recent decades, great advances have been made in sensor and monitoring technology, for example, in real-time condition monitoring of aircraft engines, as well as in off-line diagnostic testing. For systems such as military aircraft this results in large, heterogeneous datasets containing information on internal vibration, chemical composition of propellants and lubricants, corrosion, etc., as well as environmental data such as ambient temperature and humidity. The challenge for PHM is to filter and integrate such data to drive predictive models.

Comparisons Of Inferences On Software Reliability By Aggregated And Separated Recording Of Fault Counts

✦Shuen-Lin Jeng, National Cheng Kung University, No. 1, University Road, Tainan, International 701 Taiwan, ROC, *sljeng@mail.ncku.edu.tw*

Key Words: Aggregated Recording, Critical Fault-Detecting Time, Nonhomogeneous Discrete-Compound Poisson, Separated Recording, Software Reliability

In this study, we show that the separated recording of fault counts of each system in a software fault-failure processes (SFFP) is better than the aggregated recording of fault counts of all systems (used by Jeske et al., 2005). The aggregated recording approach may lose the information of variability from system to system and is difficult to identify different failure modes of SFFPs. Including the typical reliability metrics such as the average number of failures in a time interval, this study also proposes an alternative reliability index, critical fault-detecting time (CFDT) in order to provide more information for managers on judgments for software releasing times. Through simulations, our study suggests that the separated recording approach is more appropriate since it can provide more accurate estimates of CFDTs. The conventional NHPP, however, may not be suitable for some realistic practices when each failure event is caused by two or more faults. For this issue, our work considers a more general and feasible model, i.e., nonhomogeneous discrete-compound Poisson (NHDCP) model, for the software fault-failure processes.

500 Demographic Survey Sample Redesign at the U.S. Census Bureau ●

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Fifty Years Of Sample Redesign At The U.S. Census Bureau

✦James Farber, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *james.farber@census.gov*

Key Words: demographic surveys, survey methdology, history

Sample Redesign is the program that designs, selects, and disseminates updated samples for major demographic surveys conducted by the U.S. Census Bureau. The surveys and redesign have a long history, fifty years since the first redesign was conducted following the 1960 Census. This paper highlights some of the key changes in the surveys and the redesigns over that time. Several major innovations in the current redesign are also highlighted, including the use of the Master Address File for the sampling frame and the use of administrative records in statistical research. This paper also looks at some of the challenges facing demographic surveys, and the research and innovations of other organizations to respond to those challenges. Innovative methods such as responsive design may be implemented in future redesigns of Census Bureau surveys.

Frame Improvement For The Demographic Surveys Sample Redesign

✦Xijian Liu, US Census Bureau, , *xijian.liu@census.gov*

Key Words: coverage improvement, Master Address File, address base sampling

This paper will summarize the research conducted by the U.S. Census Bureau to support the development of a frame improvement system for Title-13 household surveys in the 2010 Demographic Surveys Sample Redesign. Previous coverage evaluations conducted by the U.S. Census Bureau concluded that a household survey frame developed using the Master Address File as the sole source can provide coverage comparable to that of a four-frame system currently used these surveys at the national level and that there would be risk of non-trivial coverage bias to survey estimates in certain states due to MAF undercoverage in rural areas. The U.S. Census Bureau conducted subsequent research to identify the areas that coverage improvement is needed. This paper will discuss this research and the proposed frame improvement system for the Title-13 household surveys that is based on the research results.

Sample Design Research In The 2010 Sample Redesign

✦T. Trang Nguyen, U.S. Census Bureau, 4600 Silver Hill Rd, Washington, DC 20233, *thuy.trang.ta.nguyen@census.gov*; Arielle Gerstein, United States Census Bureau

Key Words: Demographic surveys sample redesign, household surveys, survey coodination, annual sampling, multistage sample design

The sample design of the household surveys is a two stage design. In the first stage, primary sampling units (PSUs) are defined, stratified, and selected. PSUs are made up of counties or groups of contiguous counties. In the second stage, address records within each PSU are sorted by geographic and demographic information and selected systematically survey-by-survey with previously selected samples removed and the sampling interval adjusted in the next survey. In this paper, we provide an overview of the sample design for the household surveys. We also highlight research and results on major methodological changes for the 2010 sample design, including the frequency of sampling, and the use of data from the American Community Survey and administrative records in sampling.

How The U.S. Census Bureau'S Redesign Of Household Survey Sampling Impacts Field And Clerical Operations

✦Alex Winstead, United States Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746, *alexander.c.winstead@census.gov*; Rose Byrne, United States Census Bureau; Scott M. Harclerode, U.S. Census Bureau; Chris Kuwik, U.S. Census Bureau

Key Words: sample redesign, locating sample units, household surveys, field operations, group quarters, survey sample

Every 10 years the US Census Bureau uses the results from the Decennial Census to update the sample used for many of the household surveys, including CPS, SIPP, NCVS, NHIS, CE and AHS. Sample Redesigns provide an opportunity to improve operational instructions, improve coverage, and implement new technology. This paper discusses the improvements and changes to the existing processes used by the Regional Offices, Headquarters support staff, and the clerical operations supported by the US Census' National Processing Center. As a result of the 2010 Address Canvassing, a majority of household addresses in the country now have GPS coordinates. Additionally, the last decade has provided near universal access to technological tools such as online mapping software and personal navigation devices. Combining these advances can radically change how the Regional Offices locate difficult addresses. The introduction of operations to mitigate possible undercoverage (particularly for new construction group quarters) will also be discussed.

501 Human Rights

Social Statistics Section, Statistics Without Borders, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Application And Interpretation Of Multiple Systems Estimation Methods In Human Rights Research

✦Megan Price, Benetech, 480 S. California Ave, Suite 201, Palo Alto, CA 94306-1609, *meganp@benetech.org*; Daniel Manrique, Duke University; Anita Gohdes, Humboldt University

Key Words: multiple systems estimation, capture-recapture, human rights

Multiple systems estimation (MSE) methods are often introduced with the classic two-system estimator, using a motivating example from ecology. This estimator relies on four strong assumptions, which are typically then explored and adjusted for with the generalization to three or more systems and a variety of MSE methods and estimators. We find that the implications of violating these four classic assumptions, in particular the homogeneity and independence assumptions, do not translate clearly to the more-than-three-systems case. We use case studies from human rights research to present the subtle ways that misunderstanding these assumptions can lead to misapplication and misinterpretation of MSE methods. Additionally, we examine how MSE methods rely on patterns of inclusion or capture to represent the underlying population of interest even when the individual samples themselves are not representative.

Capture-Recapture Estimation Of War Deaths: Foundations, Assumptions, And Challenges

✦Nicholas P Jewell, University of California, Berkeley, 101 Haviland Hall, University of California, Berkeley, Berkeley, CA 94720, *jewell@berkeley.edu*; Michael Spagat, Royal Holloway College; Britta Jewell, Oxford University

Key Words: Capture-recapture, Multiple systems estimation, log-linear models, counting casualties

Capture-recapture estimation has been used to count elusive wildlife and human populations since the late nineteenth century. The ideas were subsequently adapted for epidemiological and social science applications, often being used to count elusive human communities. The specific application of capture-recapture estimation to casualty counting in conflict settings is a more recent development that has now been employed in sufficient examples for the methods to be taken seriously and critiqued. Here we briefly consider some questions underlying this kind of application including: When, and under what conditions, can capture-recapture estimation provide accurate and useful counts of the number of conflict casualties? What is the minimum level of data validity that is required for the methods to be applicable? How do we best choose a model for capture-recapture estimation? How do we best choose a model for capture-recapture estimation? How do we best use multiple lists? How are the results and assumptions of a particular capture-recapture analysis best translated for political consumption and the media without undermining the credibility of the approach?

Practical And Institutional Barriers To Statistical Best Practices In Political Science

✦Amelia Hoover Green, Benetech Human Rights Data Analysis Group/Yale University, 670 Apgar St., Oakland, CA 94609, *AmeliaHooverGreen@gmail.com*

Political science research on wartime violence against civilians frequently relies on quantitative data. Yet it has seldom grappled with the measurement issues afflicting political violence data. This inattention to facts on the ground may lead to serious inferential errors. However, both practical and institutional barriers have hindered the adoption of statistical best practices in conflict studies. One such 'best practice' is multiple systems estimation (MSE). In this paper, I describe a multi-method study of violence against civilians during civil war in El Salvador, which employs MSE to correct raw data on violence. I review the project's substantive findings, in particular the estimated death toll and the asymmetry of violence between government and insurgents. Additionally, I describe traditional analytical strategies for this type of research question (linear regression or qualitative analysis), and how they differ from MSE-based strategies, both methodologically and substantively. Most importantly, I consider the analysis and academic reception of this study as an example of the practical and institutional barriers to statistical best practices in political science.

Multiple Record Linkage: Generalizing The Fellegi-Sunter Theory

✦Mauricio Sadinle, Carnegie Mellon University, Pittsburgh, PA , *msadinle@stat.cmu.edu*

Key Words: EM algorithm, Fellegi-Sunter theory, Mixture models, Undirected graphs

A probabilistic method to link multiple data files is presented. This task is not trivial when there is no unique identifier available for the individuals recorded. In order to perform a record linkage for more than two data files, some authors have used record linkages for each pair of data files. However, this approach leads to discrepancies because the transitivity of the linkage decisions is not guaranteed. The method presented here is a generalization of the Fellegi-Sunter theory for record linkage for two data files. This method incorporates the transitivity of agreement since the computation of the data used to model matching probabilities. We firstly present the K-ary product of the data files to be linked as the union of all the possible subsets that contain the possible patters of agreement of the record K-tuples. Comparison data are used to model matching probabilities, incorporating the transitivity of the comparison. Also a mixture model is used to fit matching probabilities and the adjustment of this model is made through the EM algorithm. We present a method to decide the record K-tuples membership to the subsets of patterns of agreement.

Multiple Recapture Estimation In The Presence Of Unobserved Heterogeneity Using A Soft-Classification Approach

✦Daniel Manrique-Vallier, Duke University, 2800 Croasdaile Drive Ap M5, Durham, NC 27705, *dmanriq@gmail.com*; Stephen E Fienberg, Carnegie Mellon University

Key Words: Capture-Recapture, Heterogeneity, Dependence, Hierarchical Bayes, Multiple Systems Estimation, Grade of Membership

We revisit the closed population multiple recapture problem, modeling individual-level heterogeneity using the Grade of Membership model. This strategy allows us to account for the unobserved heterogeneity present in the population by postulating the existence of homogeneous latent ideal classes, and constructing a soft clustering of the individuals, where each one is allowed partial or mixed membership in all of the classes.

502 Statistical planning and data analysis in medical device trials ■

Biopharmaceutical Section, ENAR, Section on Health Policy Statistics, Section for Statistical Programmers and Analysts

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Evaluating Time Clustering Of Non-Sustained Arrythmias With The Fractal Correlation Dimension

✦Milan Seth, Boston Scientific Corporation (BSCI), Minneapolis, MN , *milan.seth@bsci.com*

Key Words: Fractal, Chaos, Competing risks, ICD inappropriate therapy

The Fractal correlation dimension (FCD) has been used in a variety of disciplines to evaluate the extent to which intermittent events (for example earthquakes) occur following a chaotic process resulting in clusters of events over time. In this presentation, we describe an application of this approach to non-sustained and diverted ventricular arrythmias recorded in patients with implanted cardiac defibrilators (ICDs) and transmitting regularly to an electronic remote monitoring system. The extent of clustering as reflected by the FCD for each sampled patient (n = 1,695) was estimated using a modified Takens estimator (calculation to be described). In order to assess evidence of overall clustering, a paired non-parametric (signed rank) test was performed comparing the FCD estimates for each patient with those of matched simulated patients with an equal number of episodes randomly distributed over an equivalent follow up period. Clustering was then assessed as a predictor of the time to first inappropriate delivery of ICD shock therapy in a proportional subdistributions hazards (Fine and Gray) model adjusting for the competing risks of death and appropriate ICD therapy.

Statistical Analysis Of A Device Trial With Excessive Numbers Of Measurements Per Subject

✦Xiaolong Shih, Boston Scientific Corporation, 4100 Hamline Ave N, South Town Square, MS 9-315, St Paul, MN 55112 USA, *alex.shih@bsci.com*; Joseph bero, Boston Scientific Corporation

Key Words: device trials, repeated measurements, generalized estimating equations

In a three-arm randomized device trial, there were excessive numbers of measurements on the primary endpoint per subject, ranging from 60 to 500. This poses extra challenges on interpreting, presenting, and analyzing the data. Several approaches will be explored and the results will be compared.

Exploratory Of Three-Way Propensity Score Matching In Drug Eluting Stent Trials

✦Hong Wang, Boston Scientific, 100 Boston Scientific Way, Marlborough, MA 01752, *Hong.Wang@bsci.com*; Songtao Jiang, Boston Scientific

Key Words: Drug eluting stent, Propensity score, Clinical trial, Matching

The propensity score matching method has been widely used in epidemiology. Recently the Greedy 1:1 matching algorithm was applied to drug eluting stent trials to assess the clinical outcome with balanced baseline characteristics for two stent treatments. We will use the approach to assess the matching for three stent groups to compare clinical outcomes in drug eluting stent trials. The drug eluting stent trial data will be used to demonstrate the difference of the two-way and three-way approaches.

Can We Have A Self-Reliant Sample Size Modification In Clinical Trials?

✦Jin Wang, Abbott Vascular, 3200 lakeside drive, santa clara, CA 95054, *jin.wang@av.abbott.com*

Key Words: adaptive design, sampling, sample size, type I error

In clinical trial methodology development, recent years have seen many statistical techniques developed so that mid-course sample size modification based on the interim treatment effect can be carried out without inflating the type I errors. However, logistic challenges still exist to implement these modifications due to the planning for the new enrollment and its associated bias concerns. In this paper, an attempt has been made to explore the re-sampling approaches so that the need for additional sample size increase might be based on a random sample from the originally planned subject set. Simulations are conducted to understand the behavior of these approaches and some discussions are provided.

A Review Of Covariate-Adjusted Endpoint Analysis

✦Zhen Zhang, Abbott Vascular, 3200 Lakeside Drive, Santa Clara, CA 95054, *zhen.zhang@av.abbott.com*

Key Words: randomized clinical trial, covariate adjustment, endpoint, variable selection

In randomized clinical trials (RCT), it is often believed that the whole randomization process will eliminate all biases and provide an ideal situation for unconditional test of the treatment effect. However, recent literatures suggested that even in RCTs, covariate-adjusted endpoint analysis (CAEA), when used appropriately, can improve the accuracy of the estimate of the effect size and therefore boost power of the endpoint

analysis. In our work, we first provide a thorough review of recent developments in CAEA. Next, through extensive and carefully designed simulations, we will demonstrate advantages as well as disadvantages associated CAEA. Last, we will illustrate through examples on how to incorporate CAEA in trial design.

503 Statistical considerations in bioequivalence, therapeutic equivalence, bio-similarity, and non-inferiority investigations

Biopharmaceutical Section, ENAR

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Stratified Analysis Of Bioequivalence And Therapeutic Equivalence Trials Data When The Metric Is Ratio Of Parameters

✦Kallappa M. Koti, Food and Drug Administration, 10903 New Hampshire Avenue, Room 3535, WO Bidg. #21, Silver Spring, MD 20993, *kallappa.koti@fda.hhs.gov*

Key Words: Non-inferiority, non-unity ratio hypothesis, Fieller-Hinkley Distribution, IUT test, generic medicines, bio-similar products.

Testing for equivalence in terms of the ratio of the parameters is often preferred rather than the difference. Fieller's method, which is often used to prove or disprove equivalence, does not always work. We propose some simpler new methods for demonstrating bioequivalence and therapeutic equivalence when the metric is a ratio of parameters and the margin is prespecified. We focus on stratified analysis. We consider comparison of means, binomial proportions, and medians of time to event endpoints. The views expressed in this presentation do not necessarily reflect those of the Food and Drug Administration.

Evaluating Bioequivalence of Two AUCs Obtained from Sparse or Incomplete Sampling in Parallel Designs

✦Steven Hua, Pfizer, Inc., , *Steven.Hua@pfizer.com*; Jihao Zhou, Pfizer, Inc.

Key Words: Bioequivalence, Area under the curve, sparse sampling, pharmacokinetics, clinical trials

Pharmacokinetic (PK) parameter such as AUC is frequently assessed in establishing bioequivalence (BE) of two drugs. For many PK studies, the sparse or incomplete sampling technique is used to collect data. In sparse sampling, subjects are independently assigned to one of the two drugs and each subject can be sampled only once. For incomplete sampling, subjects are sampled more than once but not at all timepoints. Although confidence interval approaches have been given greater attention, we show that decision rules can be established using equivalence test procedures for either the ratio or difference (in logarithm) of two AUCs. To control the type I error rate, the involved degrees-of-freedom (df) is adjusted using the Satterthwaite approximation for sparse sampling data, although Kenward-Roger estimate is preferable for incomplete sampling data. For incomplete sampling data, a linear mixed-effect model with repeated-measures data can be used to generate estimates for BE assessment. We also discuss power functions for sparse sampling designs. Depending on the structure of variances, unequal allocation of subjects could improve results relative to equal allocation.

Assessing Noninferiority In Multiregion Trial Using The Bayesian Approach

✦SATRAJIT ROYCHOUDHURY, Novartis Pharmaceutical Company, 180 PARK AVENUE, BLDG# 104 ROOM# 2G23, FLORHAM PARK, NJ 07932 United States, *satrir@gmail.com*

Key Words: multi-regional clinical trial, heterogeneity, Hierarchical bayesian model, Dirichlet process

In multi-regional clinical trials one of the key issue is how to assess the consistency of treatment effect across different regions. Recently most confirmatory trials recruit patients across the world. It is very important for regulatory authorities to ensure that the drug is effective overall. This information is also valuable for sponsors as it gives a good picture of disease and population characteristics in different regions which is valuable to plan future trials. In this article, we proposed a semiparametric bayesian methodology to assess the overall consistency of treatment effect across all regions in Phase III noninferiority trial. In particular a shared parameter model between the two endpoint are developed by assuming a semiparametric Dirichlet process prior for the shared parameter. The model will be assessed using simulation study and data analysis.

On Analysis Of The Difference Of Two Exposure-Adjusted Poisson Rate

✦Ivan S.F. Chan, Merck Research Laboratories, , *ivan_chan@merck.com*; William W.B. Wang, Merck Research Laboratories - MSD Shanghai

Key Words: Approximate exact method, Exact test, Poisson rate, Stratified analysis, Risk difference

Exposure adjusted incidence rates are often used to account for differential follow-up in evaluating treatment effects from long-term studies. For rare diseases, the Poisson distribution is a reasonable approximation for the disease occurrence. The difference of rates is sometimes used to quantify the treatment's public health impact because the reciprocal of this difference can be interpreted as "the number needed to treat (or number needed to vaccinate) in order to cure (or prevent) 1 case of disease". In this paper we focus on the analysis of two exposure-adjusted rates for both superiority and noninferiority hypothesis testing. We will present several asymptotic and exact methods for analyzing the difference of two Poisson rates with stratification. Performance of these methods will be compared in terms of type I error control and power. An example is used to illustrate these methods. Finally, we will make some general practical recommendations.

Unifying Frequentist And Bayesian Approaches In Non-Inferiority Testing

✦Mark A Gamalo, FDA/CDER/OTS/OB, 10903 New Hampshire Ave, Bldg 21 rm 3605, Silver Spring, MD 20993 USA, *mark.gamalo@fda.hhs.gov*; Ram C Tiwari, FDA/CDER/OTS/OB

Key Words: Non-inferiority, Bayesian, Frequentist, Generalized P-values

Regulatory framework recommends that novel statistical methodology for analyzing trial results parallels the Frequentist strategy, e.g., the new method must protect type-I error and arrive at a similar conclusion. Keeping these in mind, we construct a Bayesian approach for non-inferiority trials with normal response. A non-informative prior is assumed for the mean response of the experimental treatment and Jeffrey's prior for its corresponding variance when it is unknown. The posteriors of the mean response and variance of the treatment in historical trials are then assumed as priors for its corresponding parameters in the current trial, where that treatment serves as the active control. From these priors a Bayesian decision criterion is derived to determine whether the experimental treatment is non-inferior to the active control. We will also present a generalized p-value approach for assessing noninferiority that gives exact tail probabilities and is especially useful when variances are not equal between treatments. This generalized p-value approach is has a similar Bayesian construction but is entirely Frequentist. We will compare characteristics of the three approaches with

504 Advances in Copula Modeling ●

Business and Economic Statistics Section, International Indian Statistical Association

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Pair Copula Constructions Of Multivariate Copulas And Their Applications

✦Claudia Czado, Technische Universitaet Muenchen, , *cczado@ma.tum.de*

Key Words: multivariate copulas, pair copula constructions, vines, dependence, MCMC, multivariate time series

Pair copula constructions (PCC) allow to construct very flexible multivariate distributions. These are characterized by a sequence of trees called a vine structure, bi-variate copula families and marginal distributions (see Kurowicka and Cooke (2006, Wiley\& Sons) and Kurowicka and Joe (2011, World Scientific)). Two studied subclasses are C- and D-vines. The multivariate Gauss and t-distribution are special cases. This class is very useful for modeling of multivariate data in economics and finance, since it can capture non symmetric and different tail dependencies (see Aas et.al (2009, IME) and Fischer at el (2010, QF)). I will introduce PCC models and discuss different estimating methods including maximum likelihood and Bayesian inference (see Min and Czado (2010, JFES)). Since the class of PCC models is very large model selection is vital. In Min und Czado (2010, CJS) and Smith et. al (2010, JASA) two MCMC methods are developed to detect conditional independences in a given D-vine, which I will discuss. Finally I will introduce a first approach to find appropriate regular vines structures. Methods will be illustrated by applications to financial times series.

Pair Copula Constructions For Discrete Data

✦Anastasios Nicholas Panagiotelis, Technische Universitaet Muenchen, Haidhauserstrasse 1, Muenchen, International 81675 Germany, *panagiot@ma.tum.de*; Claudia Czado, Technische Universitaet Muenchen; Harry Joe, University of British Columbia

Key Words: Copulas, Discrete Data, Vine Pair Copula Constructions

Copulas provide a framework for constructing multivariate models with a wide range of dependence features. Unfortunately, applications of copula models to discrete data have been limited to date since the number of times a copula must be evaluated to compute the probability mass function (pmf) grows exponentially with the dimension of the model. Our contribution is to extend vine Pair Copula Constructions (PCCs) to the discrete case. This framework has two major advantages. First, a large degree of flexibility can be achieved by selecting different copula families to be used in the PCC. Secondly, the number of evaluations of copula functions required to compute the pmf only grows quadratically with dimension. Consequently, maximum likelihood estimation is computationally feasible even in high dimensions. We demonstrate the high quality of maximum likelihood estimates and bootstrapped confidence intervals under a simulated setting. We also illustrate the inferential potential of our model in two real data applications. We also address important model selection issues and outline interesting new directions for future research in the modeling of multivariate discrete data.

Estimation Of Copula Models With Discrete Margins

✦Michael Stanley Smith , Melbourne Business School, University of Melbourne, 200 Leicester Street, Carlton, International Australia, *mike.smith@mbs.edu*; Mohamad Khaled, University of Sydney

Key Words: D-Vine, Copula Model, Discrete Time Series;, Data Augmentation

Estimation of copula models with discrete margins is known to be difficult beyond the bivariate case. We show how this can be achieved by augmenting the likelihood with uniform latent variables, and computing inference using the resulting augmented posterior. To evaluate this we propose two efficient Markov chain Monte Carlo sampling schemes. One generates the latent variables as a block using a Metropolis-Hasting step with a proposal that is close to its target distribution. Our method applies to all parametric copulas where the conditional copula functions can be evaluated, not just elliptical copulas as in previous Bayesian work. To demonstrate the potential in higher dimensions we estimate 16 dimensional D-vine copulas for a longitudinal model of usage of a bicycle path in the city of Melbourne, Australia. The estimates reveal an interesting serial dependence structure that can be represented in a parsimonious fashion using Bayesian selection of independence pair-copula components.

Flexible Copula Models As Effective Multivariate Density Estimators

✦Robert Jacob Kohn, University of New South Wales, Australian School of Business, School of Economics, Sydney 2052, Sydney, International 2052 Australia, *r.kohn@unsw.edu.au*

Key Words: shrinkage, mixtures

Standard copula models such as the Gaussian and t copulas have proven very effective as multivariate density estimators. This paper Adds flexibility to these in a number of ways. First we consider covariance shrinkage to allow for the efficient estimation of high dimensional models. Second, we consider mixtures of copula models to more effectively capture multimodality and heavy tails. The methodology is illustrated with a number of simulated and real example.

Truncated And Simplified Regular Vines For High-Dimensional Financial Risk Modeling

✦Eike Christian Brechmann, Lehrstuhl f¸r Mathematische Statistik, Technische Universitaet Muenchen, Parkring 13, Garching, 85748 Germany, *brechmann@ma.tum.de*; Claudia Czado, Technische Universitaet Muenchen; Kjersti Aas, Norwegian Computing Center

Key Words: multivariate copula, regular vines, truncated vines, simplified vines

Introduced by Bedford and Cooke (2001, Ann. Math. Artif. Intell.) and discussed in detail in Kurowicka and Cooke (2006, Wiley & Sons) regular vines (R-vines) are a flexible class of high-dimensional dependency models which use only bivariate copulas as building blocks. Each copula can be chosen arbitrarily and the full model exhibit complex dependence patterns such as asymmetry and tail dependence. The flexibility however comes along with a strongly increasing model complexity in higher dimensions. We therefore propose innovative efficient ways of constructing R-vines. This involves the choice of R-vine trees, selection of copula types and, in particular, possible simplifications after modeling a certain number of dependency levels with copulas. Regarding the latter issue we propose different new approaches using iterative Vuong tests, AIC and BIC as well as goodness-of-fit tests. All approaches are validated in extensive simulation studies and a substantial application to the returns of the Euro Stoxx 50 members is presented. In particular, we construct an extended version of the classical CAPM, the so-called Regular Vine Market Sector Model, which is based on a simplified R-vine.

505 Bayesian spatial models

Section on Bayesian Statistical Science

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Multivariate Spatial Factor Analysis With Missingness Using Gaussian Predictive Processes

✦Qian Ren, Univerity of Minnesota, 13800 Chestnut Dr., Apt 324, Eden Prairie, MN 55344, *renxx014@umn.edu*

Key Words: Bayesian inference, Factor analysis, Gaussian predictive process, Latent factor models, Multivariate spatial process

Multivariate spatial data often arise in the natural and environmental sciences, where inferential requirements entail joint modeling of several correlated outcome variables. Hierarchical factor analysis (FA) models posit that a smaller set of latent variables can capture multivariate dependencies, thereby reducing the model dimension. In the spatial context, dimension reduction is also required with respect to the number of observed locations. Here, we demonstrate how a dimension-reducing low-rank spatial process (called a predictive process) leads to a class of computationally feasible spatial factor analysis model, thereby reducing the computational burden. A Markov chain Monte Carlo (MCMC) algorithm was developed for estimation with an emphasis toward missing data. We present sampling-based methods that condition on the observed data and recover the full posterior distribution of the missing values (along with model parameters) in a Bayesian predictive framework. Various additional modeling and implementation issues are discussed as well and we illustrate our methodology with simulated data as well as an environmental data set.

Student Paper Competitions: A Bayesian Hierarchical Model For Cosmological Parameter Inference From Supernovae Type Ia Data

✦Marisa Cristina March, Imperial College London, Astrophysics Group, Blackett Laboratory, London, International SW7 2EH UK, *marisa.march06@imperial.ac.uk*; Roberto Trotta, Imperial College London; Pietro Berkes, Brandeis University; Glenn Starkman, Case Western Reserve University; Pascal Mark Vaudrevange, DESY

Key Words: Bayes, Cosmology, Inference, Supernova Ia, Dark Energy

Supernovae type Ia (SNIa) are used to measure the apparent late time acceleration of our Universe, a phenomena often attributed to the presence of `dark energy'. The make up of the Universe is characterized by the cosmological parameters which include the total matter density, the dark energy density and the dark energy equation of state. We present a new methodology based on a Bayesian hierarchical model to extract constraints on cosmological parameters from SNIa data obtained with the SALT II lightcurve fitter. We demonstrate with simulated data sets that our method delivers considerably tighter statistical constraints on the cosmological parameters and that it outperforms the usual chisquare approach 2/3 of the times. As a further benefit of our methodology, a full posterior probability distribution for the intrinsic dispersion of the supernovae is obtained. We apply this method to recent SNIa data and find that it improves statistical constraints on cosmological parameters by about 40% w.r.t. the standard method. We constrain the intrinsic dispersion of the SNIa population, obtaining sigma intrinsic = 0.13 \pm 0.01.

Student Paper Competitions

✦Francisco M Beltran, UC Santa Cruz, 548 Seabright Ave., Apt B, Santa Cruz, CA 95062 United States, *beltran@soe.ucsc.edu*

Key Words: Global Climate Model Ensembles, Empirical Orthogonal Functions, North Pacific Sea Surface Temperature, Dynamic Linear Models

The goal of this work is to develop a general methodology to obtain joint projections of climate indexes, based on ensembles of global climate model (GCM) output and historical records. We consider Sea Surface Temperature (SST) in the North Pacific Ocean and use two ensembles of 17 different GCM results, made available in the 4th Assessment Report of the Intergovernmental Panel on Climate Change: one corresponds to 20th century forcing conditions, the other corresponds to the A1B emissions scenario for the 21st century. Given a representation of the SST spatio-temporal fields based on a common set of empirical orthogonal functions (EOFs), we use a hierarchical Bayesian model for the EOF coefficients to estimate a baseline and a set of model discrepancies. The model enables us to extract relevant temporal patterns of variability. This is used to obtain unified 21st century forecasts of relevant oceanic indexes as well as fields of forecast North Pacific

SST. We compare the forecast index for different time scales and compare the SST reconstructions to the GCMs for the 21st century. We show that finer time scales produce results that are similar to ones obtained at coarser scales.

Student Paper Competitions

✦Nilotpal Sanyal, University of Missouri - Columbia, 146 Middlebush Hall, Columbia, MO 65211-6100, *ns5x2@mail.missouri.edu*

Key Words: Bayesian, Wavelet, fMRI, Spatiotemporal analysis, Smoothing

We develop methodology for Bayesian hierarchical multi-subject multiscale analysis of functional Magnetic Resonance Imaging (fMRI) data. After modeling the brain images temporally with a standard general linear model, we transform the resulting estimated standardized regression coefficient maps by a discrete wavelet transformation to obtain a sparse representation in the wavelet space. Next, we assign to the wavelet coefficients a prior that is a mixture of a point mass at zero and a Gaussian noise. Further, we assign for the mixture probabilities a prior that depends on few hyperparameters. We develop empirical Bayes methodology to estimate the hyperparameters and use these estimates to perform inference in the wavelet space. Finally, we obtain smoothed images of the regression coefficients by inverse wavelet transformation of the posterior means of the wavelet coefficients. An application to synthetic data shows that, when compared to single subject analysis, our multi-subject methodology performs better in terms of mean squared error overall and also individually. Finally, we illustrate the utility and flexibility of our methodology with an application to a real fMRI dataset.

Bayesian Inference For Sensitivity Analysis Of Computer Simulators, With An Application To Radiative Transfer Models

✦Marian Farah, University of California, Santa Cruz, 111 Echo Street, Santa Cruz, CA 95060, *marian@ams.ucsc.edu*; Athanasios Kottas, University of California, Santa Cruz

Key Words: Gaussian process prior, Sensitivity analysis, Leaf-Canopy Model, Main effects, Sensitivity indices

Sensitivity analysis of computer simulators is a useful tool for quantifying the impact of uncertainty in the computer simulator inputs on the computed output. The influence of each input and how uncertainty in the output is apportioned amongst the inputs are determined by calculating the main effects and sensitivity indices of the computer simulator inputs. Typically, these quantities are computed using Monte Carlo methods, which require a large number of computer simulator runs, making the calculations infeasible if the simulator is computationally expensive. Bayesian methods have been used to tackle sensitivity analysis of computationally expensive simulators through building a statistical emulator for the computer simulator output, typically, based on a Gaussian process (GP) prior for the simulator output function. We develop an approach that unifies GP-based global sensitivity analysis tools and extends semi-Bayesian approaches to a fully Bayesian methodology. The approach is utilized to carry out sensitivity analysis of the Leaf-Canopy Model, a radiative transfer model that simulates the interaction of sunlight with vegetation.

506 Opportunities and Challenges in Statistical Work with Administrative Record Data ■●

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

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

A Case Study In Using Household Information From Administrative Databases

✦Martin Pantel, Statistics Canada, 110 Tunney's Pasture Driveway, R.H. Coats bldg, 17-L, Ottawa, ON K1A 0T6 Canada, *Martin.Pantel@statcan.gc.ca*

Key Words: Frames, Administrative data, Paradata

The Canadian Child Tax Benefit (CCTB) is a non-taxable amount paid monthly to help eligible Canadian families raise children until the age of 18. The list of applicants to the CCTB was used as a sampling frame for the first time in 2010, for Statistics Canada's Survey of Young Canadians (SYC). Data collection for this new cross-sectional survey, which includes both telephone interviews and personal visits, took place from November 2010 to February 2011. The presentation will briefly describe how the sampling frame was built from the CCTB database, and assess the quality of the contact information (names, birthdates, phone numbers and addresses) by presenting results of a small-scale test conducted in February 2010, as well as early results from the SYC data collection activities. The focus will be on reporting key collection paradata and survey results, and comparing some of these with results from projects using other types of frames. We will assess the accuracy of the frame information, patterns of non-response, and the success of tracing activities based on the multiple leads made available by the SYC frame.

Sources Of Revision In A Multi-Mode Survey

✦Keven Bosa, Statistics Canada, 100 Tunney's Pasture Driveway, RH Coats, 11-D, Ottawa, ON K1A 0T6 Canada, *keven.bosa@statcan.gc.ca*

Key Words: multi-mode, administrative data, imputation, revision

Statistics Canada's Quarterly Industry Revenue Indices (QIRI) provides sub-annual indicators of economic activity for 13 selected business and consumer services. This program consists of a census where administrative and survey data are combined to measure changes in the quarterly operating revenues. For simple structured enterprises, data are extracted from an administrative file. A questionnaire is used to collect information for enterprises with complex operational structures. A preliminary index is produced three months after the end of the reference quarter. A revision of this index is done three months later using more up to date information. Investigations have shown that revised indices are usually, although not systematically, lower than preliminary numbers. It is known that some enterprises for which administrative data are used in preliminary indices are later identified as inactive (dead) by the time the indices are revised. Although this represents an important source of revisions, a study was done to quantify the various sources of revision and verify the overall impact of the inactivation issue on revision rates. Details and results of the study will be presented.

Utilizing Automated Statistical Edit Changes in Significance Editing

✦Wendy Barboza, National Agricultural Statistics Service, 3251 Old Lee Hwy, Room 305, Fairfax, VA 22030, *wendy_barboza@nass.usda.gov*; Kay Turner, National Agricultural Statistics Service

Key Words: significance editing, statistical data editing, selective editing, outlier detection, NASS

The National Agricultural Statistics Service (NASS) is a statistical agency within the U.S. Department of Agriculture (USDA) that conducts hundreds of surveys every year and prepares reports covering virtually every facet of U.S. agriculture. NASS' traditional approach has been to perform a manual edit review of all questionnaires for most surveys. As staff resources become more constrained, the agency has embraced technological advances. The goal of significance editing, defined as statistical data editing; selective editing; and outlier detection, is to (1) reduce the time and effort spent manually reviewing/correcting survey questionnaires, without damaging the quality of the resulting data, and (2) focus the manual effort on the accuracy of the survey questionnaires that strongly impact the overall results. During the survey process, the most influential records are identified by calculating a unit-level score based on the changes made by the automated statistical data edit. This paper provides details on these unit scores as well as the implementation of the significance editing concepts.

Trends In Record Linkage Refusal Rates: Characteristics Of National Health Interview Survey Participants Who Refuse Record Linkage

✦Donna M Miller, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *DMMiller1@cdc.gov*; Renee M Gindi, National Center for Health Statistics; Jennifer D Parker, National Center for Health Statistics

Key Words: administrative data, refusal to link, health survey

The National Health Interview Survey (NHIS) is a principal source of information on health in the United States. NHIS data are linked to administrative records from several federal agencies to maximize the scientific value of the survey. Records are linked by matching administrative data with personally identifiable information (i.e., name, date of birth, Social Security Number (SSN), and Medicare health insurance claim number (HIC)) collected from NHIS participants. Participants who refused to provide SSN and HIC were implicitly considered to have refused record linkage. This refusal rate has increased over time, reducing the number of NHIS participants eligible for record linkage. Beginning in 2007, NHIS attempted to decrease linkage refusal rates by: requesting only the last four digits of SSN and HIC numbers (instead of all nine); adding a short introduction before asking for SSN; and asking participants for their explicit permission to link a record even if SSN or HIC had not been provided. We will examine the impact of these changes on record linkage refusal rates, and explore the additional influence of participant sociodemographic and health characteristics on refusal.

507 Adapting a Randomized-Based Perspective in Your Introductory Statistics Course: Roadblocks, Risks and Rewards

Section on Statistical Education

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Adopting A Randomization-Based Perspective In Your Introductory Statistics Course: Roadblocks, Risks, And Rewards

✦Rachel Mary Dunwell, Rhodes College, 2000 N. Parkway, Memphis, TN 38112, *dunwellr@rhodes.edu*; ✦Alan Rossman, California Polytechnic State University - San Luis Obispo, Department of Statistics - Cal Poly, San Luis Obispo, CA 93407, *arossman@calpoly.edu*; ✦Patti Frazer Lock, St. Lawrence University, 23 Romoda Drive, Canton, NY 13617, *plock@stlawu.edu*; ✦Nathan Tintle, Hope College, ; ✦Megan Spencer, Rhodes College,

Key Words: bootstrapping, permutation testing, resampling, introductory statistics, pedagogy, randomization techniques

In some introductory statistics courses, the emphasis placed on traditional normality-based hypothesis testing and confidence interval construction is increasingly giving way to randomization-based analogues. The introductory statistics courses at a growing number of institutions, including Rhodes College, St. Lawrence University, and Cal Poly - San Luis Obispo, now feature randomization-based methods to some degree, and sometimes to a very large degree. Using tools like R and Fathom, students in such courses quickly bypass formulas and tables, and rely instead on techniques that allow greater visualization and intuition to conduct statistical investigations. Panelists at this session will provide a brief overview of how (and to what extent) they have incorporated randomization-based methods into their introductory statistics courses, and what the results have been. After addressing some obvious concerns like "How will students react when all the p-values are different?" and "Can your students still do a basic t test?" and "How will my colleagues in the biology department react?", the panel will open the floor for discussion.

508 Survival Analysis in the Presence of Complex Censoring

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

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Buckley-James-Type Estimator With Right-Censored And Length-Biased Data

✦Jing Ning, The University of Texas School of Public Health, Division of Biostatistics,, The University of Texas School of Public Health, Houston, TX 77030, *jing.ning@uth.tmc.edu*; Jing Qin, National Institute of Allergy and Infectious Diseases; Yu Shen, The University of Texas M. D. Anderson Cancer Center

Key Words: Accelerated failure time model, Buckley-James Estimator, Estimating equation, Length-biased sampling, Prevalent cohort

We present a natural generalization of the Buckley-James-type estimator for traditional survival data to right-censored length-biased data under the accelerated failure time model. Length-biased data are often encountered in prevalent cohort studies and cancer screening trials. Informative right censoring induced by length-biased sampling creates additional challenges in modeling the effects of risk factors on the unbiased failure times for the target population. We evaluate covariate effects on the failure times of the target population under the accelerated failure time model given the observed length-biased data. We construct a Buckley-James-type estimating equation, develop an iterative computing algorithm, and establish the asymptotic properties of the estimators. We assess the finite-sample properties of the proposed estimators against the estimators obtained from the existing methods. Data from a prevalent cohort study of patients with dementia are used to illustrate the proposed methodology.

Fractional Polynomials In Analyzing Interval-Censored Time-To-Event Data

✦Din Chen, University of Rochester, P.O. Box SON, 601 Elmwood Ave, Rochester, NY 14642 USA, *din_chen@urmc.rochester.edu*; Lili Yu, Georgia Southern University; Yuhlong Lio, University of South Dakota

Key Words: Time-to-event data, Interval-censoring;, hazard function, Cox proportional hazards regression, fractional polynomials

Interval censored time-to-event data along with complete-time and right/left-censored time-to-event are generated in most oncology clinical trials especially from cancer scan within some specific time intervals. The extension of the well-known Cox regression is discussed in this talk with fractional polynomials as the approximation to the baseline hazard function. A likelihood approach is used to select the best fractional polynomial as well as estimating the model parameters with associated statistical inference for treatment effect. The application of this method is demonstrated by a simulation study and to a real breast cancer clinical trial data

Semiparametric Estimation Of Cumulative Treatment E?Ect In The Presence Of Dependent Censoring

✦Qi Gong, Department of Biostatistics, University of Michigan-Ann Arbor, 2180 St Francis Dr, Apt F142, Ann Arbor, MI 48104 USA, *gongqi@umich.edu*; Douglas E. Schaubel, University of Michigan

Key Words: cumulative hazard, stratification, dependent censoring, survival analysis, inverse weighting

In time to event data observed in medical studies, nonproportional hazards and dependent censoring are common issues when comparing group-specific mortality. The group effect on mortality may vary over time, as opposed to being constant. One remedy is adopting a parametric form to model the time-dependent pattern. However, it is generally difficult to verify the correctness of the chosen parametric function. Moreover, investigators tend to be more interested in cumulative effects on mortality (e.g., if and when the survival curves cross) rather than the instantaneous effect. Estimators are no longer consistent in the presence of dependent censoring, which may occur when both censoring and death depend on the same time dependent covariates. Therefore, we propose an estimator for the cumulative group effect on survival in the presence of nonproportional hazards and dependent censoring. The proposed estimator is based on the cumulative hazard function, assumed to follow a stratified Cox model. No functional form needs to be assumed for the nonproportionality. Asymptotic properties are derived and evaluated in simulation studies. The proposal method is applied to SRTR data.

Quantile Regression For Dependently Censored Data

✦Shuang Ji, Emory University, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322 US, *sji@emory.edu*; Limin Peng, Emory University

Key Words: quantile regression, survival analysis, dependent censoring

Quantile regression is known for its flexibility in accommodating varying covariate effects and has attracted growing interests in survival analysis. Considerable research effort has been devoted to study quantile regression under independent censoring assumptions. However, dependent censoring often occurs in practice and ignoring the association between censoring and the event time of interest may lead to substantial biases. In this work, we develop a quantile regression method that appropriately adjusts for dependent censoring when making inferences on covariate effects on marginal quantiles of the event time outcome. We propose valid estimation and inference procedures, along with an efficient and stable algorithm. We establish the uniform consistency and weak convergence of the resulting estimators. The finite-sample performance of our approach is assessed by extensive simulation studies. We illustrate the practical utility of our method via an application to a stroke clinical trial.

Weighted Logrank Tests For Interval Censored Data When Assessment Times Are Related To Treatment But Not To Event Time

✦Michael P. Fay, National Institute of Allergy and Infectious Diseases, 6700A Rockledge Dr, Room 5133, Bethesda, MD 20817, *mfay@niaid.nih.gov*; Joanna H Shih, National Cancer Institute

Key Words: interval censoring, logrank test, permutation test, small sample size, imputation

We consider weighted logrank (WLR) tests for interval censored data when assessment times may depend on treatment and for each individual we only observe two assessment times, the one before and the one after the event of interest. We review that for large samples with a small number of unique assessment times, the score test under the grouped continuous model (GCM) retains the type I error even under assessment-treatment dependence. Although the permutation form of the WLR test under the GCM is theoretically valid under assessment-treatment independence, we give heuristics and simulations to show that even under assessment-treatment dependence this type of WLR test approximately retains the type I error in all but some very extreme cases. Next, we study and extend the imputation WLR tests of Huan, Lee, Yu (2008, Stat in Med), showing that the distribution of the rank-like scores asymptotically do not depend on the assessment times. We confirm through simulation that the imputation WLR tests retain the type I error even when the number of individuals in each treatment group are small and possibly unbalanced in size and with assessment-treatment dependence.

Semiparametric Methods For The Analysis Of Failure Time Data With Outcome-Dependent Sampling And Dependent Censoring

◆Hui Zhang, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109 U.S.A., *huizh@umich.edu*; Douglas E. Schaubel, University of Michigan; John D. Kalbfleisch, Department of Biostatistics, University of Michigan

Key Words: Dependent censoring, Outcome-dependent sampling, Semiparametric methods, Survival analysis

Outcome-dependent sampling (ODS) is an efficient and cost-saving sampling scheme, wherein subjects are selected into the study based on the outcomes of interest (i.e., death, survival). Most methods for analyzing ODS-based data have an underlying assumption that subjects are censored in a manner independent of the failure rate. However, this assumption is often violated in public health studies. For example, wait-listed end-stage liver disease patients may receive a liver transplant and therefore not die on the waitlist, an issue which could produce substantial bias in the estimation of waitlist mortality if treated as independent censoring. We propose methods, based on weighted estimating equation with a double-inverse-weighting scheme which combines weights corresponding to the probability of remaining uncensored and the probability of being sampled. The proposed estimators of the regression parameter are shown to be consistent and asymptotically normal. Finite sample properties of the proposed estimators are examined through simulation studies. The proposed methods are applied to data from the Scientific Registry of Transplant Recipients.

509 High Dimensional Unsupervised and Supervised Methods

Biometrics Section, ENAR

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Simulation Of Gei And Qei Using A Genotype To Phenotype Model With Physiological Parameters Without Gei

◆Paulo C. Rodrigues, Wageningen University, International The Netherlands, *paulocanas@fct.unl.pt*

Key Words: genotype by environment interaction, QTL by environment interaction, crop growth model, multi-environment trials, simulation, pepper

A different response of genotypes across environments is frequent in multi-location trials and is known as genotype by environment interaction (GEI). When the analysis is made in the whole genome, the interactions of interest are between QTL and environment (QEI). Simulation tools such as genotype to phenotype models have proved to be useful to better understand these interactions. This paper presents a crop growth model (i.e. a non-linear function of purely genotypic components and environmental inputs which integrates over time) with 7 physiological parameters, to simulate two-way tables with yield. After modelling the genetic basis under each of the 7 model parameters we are able to predict yield from simple DNA profiles together with environmental variables. Standard statistical models, such as AMMI models, linear mixed models, and factorial regression, are used to explore and understand the GEI and QEI of the two-way tables. The results from these analyses are proved to be directly linked to the original model parameters. GEI and QEI were simulated for sweet pepper, and the QTL detected for final yield matched the QTL assigned to most of the model parameters.

Causal Models for Randomized Trials with Two Active Treatments and Continuous Compliance

◆Yan Ma, Weill Medical College of Cornell University, Hospital for Special Surgery, 10021, *yam2007@med.cornell.edu*; Jason Roy, University of Pennsylvania; Bess Marcus, Brown University

Key Words: copula model, Gauss-Hermite quadrature, Partial compliance, Potential outcomes, Principal stratification

In many clinical trials, compliance with assigned treatment could be measured on a continuous scale (e.g., the proportion of assigned treatment actually taken). In general, inference about principal causal effects can be challenging, particularly when there are two active treatments; the problem is exacerbated for continuous measures of compliance. We address this issue by first proposing a structural model for the principal effects. We then specify compliance models within each arm of the study. These marginal models are identifiable. The joint distribution of the observed and counterfactual compliance variables is assumed to follow a Gaussian copula model, which links the two marginal models and includes a dependence parameter that cannot be identified. This dependence parameter can be varied as part of a sensitivity analysis. We illustrate the methodology with an analysis of data from a smoking cessation trial. As part of the analysis, we estimate causal effects at particular levels of the compliance variables and within subpopulations that have similar compliance behavior.

Model-Based Clustering For Rna-Seq Data

◆Yaqing Si, Iowa State University, Snedecor 2418, Ames, IA 50010, *siyaqing@iastate.edu*; Peng Liu, Iowa State University

Key Words: Model-Based Clustering, RNA-Seq Data, Hierarchical Clustering, Gene Expresion, Annotation

Recently, ultra-high throughput sequencing technology has provided an unprecedented way to study gene expression. The resulting data, termed as RNA-Seq data, is ultra-high dimensional, discrete and highly skewed. Novel statistical methods are called for its analysis. Clustering is an important method to explore the gene expression patterns and has been widely applied to gene expression data generated from microarray technology. The same clustering methods used to analyze microarray data has been applied to transformed RNA-Seq data without evaluation of its performance. Here, we examine the performance of the clustering algorithms currently applied to RNA-Seq data and propose a model-based clustering algorithm for RNA-Seq data that is shown to be better in simulation study and real data analysis .

Assessing Microarray Quality

◆Matthew Nicholson McCall, University of Rochester Medical Center, Department of Biostatistics and Computational Biol, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *mccallm@gmail.com*; Peter Murakami, Johns Hopkins School of Medicine; Rafael Irizarry, Johns Hopkins School of Public Health

Key Words: microarray, quality, single-array, Affymetrix

For microarrays to become generally accepted as a reliable technology, statistical methods for assessing quality will be an indispensable component; however, there remains a lack of consensus in both defining and measuring microarray quality. We provide a precise definition of microarray quality and review existing quality metrics in light of this definition. The best-performing quality metrics require multiple arrays to be assessed simultaneously. While such metrics are adequate for bench science, as microarrays begin to be used in clinical settings, single-array metrics will be indispensable. To this end, we define a single-array version of a popular multi-array quality metric and show that this metric is comparable to the best multi-array metrics. Finally, we use this new quality metric to assess the quality of over 22,000 publicly available microarrays. We find that approximately 10% of these arrays are of poor quality. Moreover, the quality of microarray measurements varies greatly from hybridization to hybridization, study to study, and lab to lab, with some experiments producing unusable data.

An Integrative Approach To Comparing And Normalizing Gene Expression Data Generated From Rna-Seq, Microarray, And Rt-Pcr Technologies

✦Zhaonan Sun, Department of Statistics, Purdue University, Department of Statistics, Purdue University, 250N. University Street, West Lafayette, IN 47907-2066, *sunz@stat.purdue.edu*; Yu Zhu, Department of Statistics, Purdue University

Key Words: RNA-Seq, Microarray, Gene expression level, Normalization

DNA microarray and RNA-seq both provide genome-wise expression level measurements, in many comparison studies, data from the same biological cell line are generated from each platform. Usually the gene expression levels are calculated from each platform separately, then correlation coefficients across replicates is used to assess the reproducibility of each platform. We argue that in the presence of technical and biological variations, correlation coefficient can be biased. In the comparison of accuracy of these platforms, qRT-PCR data is often regarded as the gold standard, correlation coefficients with qRT-PCR indicate the accuracy, and in differential expression (DE) studies, error rates are assessed by comparing DE result to that from qRT-PCR data. Since qRT-PCR itself also suffers from variation, ignoring this will also lead to bias. In this study we propose an integrative method, we combine data from different platforms together to get a better estimation for the expression level, then base on this we make the between-platform comparison in a statistically sound way. Our method is also applicable to differential expression study.

Fdr Control For Discrete Test Statistics

✦Isaac Dialsingh, Penn State University, Department of Statistics, 331A Thomas Building, University Park, PA 16801, *iud101@psu.edu*

Key Words: Family wide error rate, False discovery rate, Multiple hypothesis testing, Discrete test statistics

In this talk I present some ways we can control the false discovery rate (FDR) when the test statistics are discrete. When the test statistics are continuous, the distribution of the p-values under the null distribution is Uniform(0,1). This has made it relatively easy to estimate the proportion of true nulls which is a key parameter in estimating and controlling the FDR. When the test statistics are discrete, the Uniform(0,1) assumption under the null hypothesis does not hold. This makes it difficult to estimate the proportion of true nulls. As well, methods that use the p-values to control FDR will be conservative, because not all p-values are achievable. This talk covers methods for estimating the proportion of true nulls and for estimating FDR when the test statistics are discrete. Examples are given using RNA-seq and SNP data.

Adaptive Thresholding At The Peptide-Level Improves Estimate Of Relative Protein Abundances From Lc-Ms Experiments

✦Yuping Zhang, Stanford University, , *yupingz@stanford.edu*; Zhengqing Ouyang, Stanford University; Wei-Jun Qian, Pacific Northwest National Laboratory; Richard D. Smith, Pacific Northwest National Laboratory; Wing Hung Wong, Stanford University; Ronald W. Davis, Stanford University

Key Words: proteomics, LC-MS, hierarchical mixed model

Quantitative comparison of the relative abundances of proteins from different biological conditions is an important aspect for LC-MS based proteomics. Because LC-MS experiments perform measurements at the peptide level, it is often challenging to obtain relative protein abundances based on these peptide level measurements due to the issues of missing data, varying numbers of peptides per protein, as well as potential multiple isoforms per protein. Here we report a method termed as protein expression through adaptive thresholding (PEAT), which utilizes a two-stage hierarchical additive mixed model for peptide-level modeling of LC-MS data, and adaptively select peptides for protein-level abundance estimation and inference. Tests on simulation data, spiking-in data and biological data from human plasma samples following burn injury demonstrate that our method provides a more accurate estimation of relative protein abundances as well as more powerful detection of differentially expressed proteins than current existing methods.

510 Bayesian Modeling in Life Sciences 2

Section on Bayesian Statistical Science

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Bayesian Inference for Nonlinear Mixed-Effects Tobit Models with Skew-Elliptical Distributions

✦Getachew Dagne, University of South Florida, 13201 Bruce B. Downs Blvd., MDC 56, Tampa, FL 33612 U.S., *gdagne@health.usf.edu*

Key Words: Tobit model, Mixed-effects models, censored data, Bayesian inference, HIV/AIDS

Censored data are characteristics of many bioassays in HIV/AIDS studies where assays may not be sensitive enough to determine gradations in viral load determination among those below a detectable threshold. Not accounting for such left-censoring appropriately can lead to biased parameter estimates in most data analysis. To properly adjust for left-censoring, this paper presents an extension of the Tobit model for fitting nonlinear dynamic mixed-effects models with skew-elliptical distributions. Such extensions allow one to specify the conditional distributions for viral load response to account for left-censoring, skewness and heaviness in the tails of the distributions of the response variable. A Bayesian modeling approach via Markov Chain Monte Carlo (MCMC) algorithm is used to estimate model parameters. The proposed methods are illustrated using real data from an HIV/AIDS study.

Combining Longitudinal Studies On Bone Mineral Content Of Infants

✦Ramu Goud Sudhagoni, South Dakota State University, Rm 218, Department of Mathematics and Statistics, Harding Hall, Brookings, SD 57007, *ramu.sudhagoni@sdstate.edu*; Gemechis Djira, South Dakota State University; Howard Wey, South Dakota State University; Bonny Specker, South Dakota State University

Key Words: Longitudinal Data, Bayesian Hierarchical Modeling, Bone Mineral Content

Bone mineral content accrual in infants may depend on several factors, namely growth, mineral intake and possibly other dietary factors (vitamin D, protein intake, etc.). Longitudinal studies that have investigated the relationship between changes in body fat and lean mass on bone accrual during growth are inconsistent. No studies have been reported that describe the influence of changes in body composition with changes in bone accrual during the first year of life. Previously, three studies investigated the influence on bone accrual due to different mineral intake (Study 1), timing of introducing solid foods to the baby (Study 2), and gross motor activity (Study 3) during the first year of life. In this research our aim is to combine three longitudinal studies using Bayesian hierarchical modeling approach to determine whether greater gains in fat and lean mass during infancy have a positive or negative influence on bone accrual.

A Bayesian Zero-Inflated Mixed Lognormal Regression Model For Emergency Department (Ed) Diversion Time

✦Rongwei Fu, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd., Portland, OR 97239 U. S. A, *fur@ohsu.edu*; Daniel A. Handel, Oregon Health & Science University

Key Words: Bayesian model, zero-inflated log-normal model, mixed model, ED diversion

Ambulance diversion occurs when one Emergency Department (ED) is too crowded and the ambulance is re-routed to a different ED. Ambulance diversion was once thought to be a “novel” solution for the problem of ED crowding. However, it is now understood that diversion is not an effective means for alleviating over-crowding. It is important to investigate predictors associated the diversion. However, the distribution of diversion time is highly skewed with many days without diversion (diversion time = 0), and diversion time of days on divert still tending to right skewed. It is important to evaluate both the decision to go on diversion and the time on diversion across multiple EDs. Therefore, we propose a Bayesian zero-inflated mixed lognormal regression model to evaluate predictors for both the decision to go on diversion and the time on diversion. The model is illustrated using data from nine EDs across US.

Bayesian Multiple Testing Under Dependence With Application To Functional Magnetic Resonance Imaging

✦D. Andrew Brown, University of Georgia Department of Statistics, 101 Cedar St, Athens, GA 30602, *d.andrew.brown@gmail.com*; Nicole Lazar, Department of Statistics; Gauri Sankar Datta, University of Georgia Department of Statistics

Key Words: multiple testing with dependence, Bayesian statistics, functional magnetic resonance imaging

In the analysis of high-throughput data, a massive number of hypotheses are tested simultaneously. Correcting for multiple testing becomes problematic because relatively simple procedures such as the Bonferroni correction are overly conservative, whereas ignoring the problem altogether leads to a very high number of false rejections. This is especially true when trying to identify sparse signals. Many other multiple testing procedures, both Bayesian and frequentist, rely on the assumption of independence of the data. One particular Bayesian procedure for the simultaneous testing of independent data was given in Scott and Berger (2006). We extend this method by introducing a conditional autoregressive (CAR) model to account for spatial dependence. The model is applied to data from a functional magnetic resonance imaging (fMRI) study and compared to results obtained by assuming independence of the data.

Assessing The Risk Of Rising Temperature On Brook Trout: A Dynamic Linear Model Approach

✦Ciro Velasco-Cruz, Virginia Tech, 1236 University City Blvd., Apt L129, Blacksburg, VA 24060 United States, *cvelasco@vt.edu*; Eric Smith, Virginia Tech

Key Words: Dynamic linear models, Spatio-Temporal, Brook Trout

Modeling the relationship between air temperature and water temperature is a critical component in the management of aquatic resources. One important application is the management of fisheries that have temperature sensitive fish. Co-located devices were used to obtain air and water temperature for summer periods from 100 locations in Viriginia that have native brook trout populations. A dynamic linear model was developed that studies the relationship between air and water temperature, allowing for spatial and temporal correlation. The model also allows the inclusion of land use, solar gain and other site level characteristics that might influence the relationship. The model is used to predict the risk to fish at sites and how that risk changes over time. The model also allows ranking of sites with regard to risk and hence can aid management in prioritizing decisions about restoration and preservation.

Bayesian Approaches For Estimating Prevalence Based On Pool Screening When Observing Zero Positive Pools

✦Thomas Birkner, University of Alabama at Birmingham, 2400 Hawksbury Lane, Hoover, AL 35226, *birkner@bellsouth.net*; Immaculada Aban, University of Alabama at Birmingham; Charles Katholi, University of Alabama at Birmingham

Key Words: Pool screening, sequential Bayes, objective prior

Pool screening is a method that combines individual units into pools. Each pool will either test positive (at least one of the units is positive) or negative (all units are negative). Pool screening is commonly applied to the study of tropical diseases where pools consist of vectors that can carry the disease (e.g. black fly). The goal is to estimate the proportion of infected vectors. When control programs are effective and the surveillance phase is in place, entomologists may encounter data with no positive pools. The traditional maximum likelihood estimate of the prevalence is zero - an unrealistic value. A sequential Bayesian approach can incorporate results from previous years and will provide a more sensible prevalence estimate even if all of the pools collected at this time are negative for the trait of interest. This paper proposes and evaluates the performance of a Bayesian approach to this problem. Through simulation, we investigate the amount of data (pool size, number of years) required such that the type of objective prior chosen (e.g. Bayes/Laplace, Jeffreys) does not make a significant difference with respect to the estimate of prevalence.

Model Assessment In Multistate Transition Models Through Joint Use Of Bayesian Variable Selection And Bayes Factors

✦David Engler, Department of Statistics, Brigham Young University, , *engler@byu.edu*; Brian Healy, Department of Neurology, Harvard Medical School

Key Words: Bayesian variable selection, Bayes factor, Multistate transition models, Model selection

In multiple sclerosis (MS), the main clinical outcome measure is an ordinal score, the expanded disability severity score (EDSS). One of the main goals of MS research is to accurately model transitions between EDSS states over time. This objective first requires a number of significant modeling decisions, including selection of an appropriate link function and whether or not to collapse EDSS states. Then, under a selected model framework, it is also expected that the probability of transitioning from one state to another will vary across subjects; modeling will depend upon identification of transition-specific covariate effects. Historically, both sets of decisions are often made in an ad hoc manner and are not formally justified. We propose joint use of Bayesian variable selection and Bayes factors in Markov transitional models to allow for both the selection of model framework and the identification of transition-specific covariate effects. Methods are assessed using both simulated data and data collected from the Partners MS Center in Boston, MA.

511 Applications in Risk Analysis

Section on Risk Analysis, Section on Statistics in Defense and National Security, Section on Statistics in Epidemiology, Section on Quality and Productivity

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Equivalence Of Improvement In Area Under Roc Curve And Linear Discriminant Analysis Coefficient Under Assumption Of Normality

✦Olga Demler, Boston University, 30 Edge Hill Rd, Newton, MA 02467 USA, *demler@bu.edu*; Michael Pencina, Boston University; Ralph B. D'Agostino, Sr, , Boston University

Key Words: AUC, ROC, risk prediction, linear disciminant analysis, logistic regression, significance of the predictor

Area under the Receiver Operating Characteristics Curve, (AUC of ROC) is a widely used measure of discrimination in risk prediction models. This study was motivated by numerous reports that often the added predictor is statistically significantly associated with the outcome but fails to produce significant improvement in the AUC. We demonstrate that under the assumption of multivariate normality and employing linear discriminant analysis to construct the risk prediction tool, statistical significance of the new predictor(s) is equivalent to the statistical significance of the increase in AUC. We extend this result for unequal, non-proportional variance-covariance matrices of predictors within cases and non-cases. The result holds asymptotically for logistic regression under assumptions of normality and equal covariance matrices. Our practical example from the Framingham Heart Study data suggests that the finding might be sensitive to the assumption of normality.

Truncated Sequential Procedures

✦Claudia Schmegner, DePaul University, 2320 N Kenmore Ave, Chicago, IL 60614, *cschmegn@condor.depaul.edu*

Key Words: Lattice, Markov Chain, Truncated Sequential Plan

Sequential planning with the restriction that the number of groups is absolutely bounded by some fixed number K is considered. These truncated sequential plans will stop sampling after at most K groups of observations are collected. Practically, it reflects the following situation. There are only K days available to complete a certain study (e.g. clinical trials, to get FDA approval or investigate conditions of the field before it is time for farming). Every day, a sample of the desired size is selected. Based on the interim analysis performed, an inference is made or another batch of observations is collected as long as the total number of groups does not exceed K. Truncated sequentially planned probability ratio tests (TSPPRTs) on a lattice are considered in detail. Construction and methods of evaluation of TSPPRTs are proposed.

Misuse Of Delong Test To Compare Aucs For Nested Models

✦Olga Demler, Boston University, 30 Edge Hill Rd, Newton, MA 02467 USA, *demler@bu.edu*; Michael Pencina, Boston University; Ralph B. D'Agostino, Sr, , Boston University

Key Words: AUC, ROC, risk prediction, discrimination, DeLong test, U-statistics

Area under the Receiver Operating Characteristics Curve, (AUC of ROC) is a widely used measure of discrimination in risk prediction models. Mann-Whitney statistics is used as a non-parametric estimator of AUC. The difference of two AUCs is often tested by DeLong test. This study was motivated by numerous reports that often the added predictor is statistically significantly associated with the outcome but fails to produce significant improvement in the AUC. We show that DeLong test can not be applied to test AUC improvement for nested models for any continuous distribution of the data and very general class of statistical models including logistic regression. First we show empirically that distribution of the difference of two AUCs from nested models is very different from the one used by the DeLong test. We use theory of U-statistics to explain this contradiction by showing that

the difference of two AUCs belongs to a degenerate class of U-statistics which has different asymptotic distribution than the one used by the DeLong test. It results in substantial (up to 60%) loss of power by the DeLong test. Possible solutions are discussed.

Estimation Of Catastrophic Losses And Tail Distributions Through Some Extreme Value Distributions

✦Fan Yang, Florida International University, FL 33199, *fyang001@fiu.edu*; Sneh Gulati, Florida International University; B. M. Golam Kibria, Florida International University

Key Words: Extreme Value, Estimation, Loss, Risk, PML, VaR

This paper will review various extreme value distributions for estimating the catastrophe losses and tail distributions. We will compare the performance of the distributions by using a goodness of fit test. Some useful extreme value distributions will be recommended for the practitioners in the field of risk analysis.

The Impact Of Previous Claims On The Future Claims: An Application To The French Automobile Insurance Market

✦Michel Grun-Rehomme, University Paris-Pantheon, 12, Place de Pantheon, Paris, 75005 FRA, *Michel.Grun-Rehomme@ensae.fr*

Key Words: Solvency, ERM, ORSA, Risk Appetite

The risk of each insured motor vehicle is not predictable and is known only retrospectively. Unlike the collective risk is predictable when we have the experience of the past later observed over a relatively large population comparable to the existing portfolio. In this study, we would like to examine empirically on data from a French insurer, if the accident record (before the reference year) and the current claims are a good predictor of future claims, provided the requirements of the class of risk (or case pricing) of the insured. It is assumed, depending on the accident record, that each class of risk has two categories of drivers insured: low-risk and high-risk. Using a negative binominal distribution and a Bayesian approach, we show that the probability of being a low risk driver is more important in the absence of losses (or with a single claim) and that conversely the probability of an insured high risk greatly increases when the insured has 2 or 3 losses during the reference year. Of course the level of probability varies according to risk classes. In all cases, the previous accidents are a good indicator of future claims.

512 The Many Faces of Kernel Methods ●

Section on Nonparametric Statistics, ENAR, Section on Statistical Computing

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Semiparametric Estimation For Mean Response With Nonresponse Data

✦Zonghui Hu, National Institutes of Health/NIAID, 6700 Rockledge, Bethesda, MD 20817 USA, *huzo@niaid.nih.gov*

Key Words: propensity score, inverse propensity weighting, mixed-effects modelling, robust estimation

We study the effect of continuous versus intermittent antiretroviral therapy on CD4 responses in HIV infected patients undergoing IL-2 cycle. Each patient in the study goes through multiple switches between continuous and intermittent IL-2 cycles, where switches are not randomized but depend on the patient's disease status. We propose a causal estimation of the treatment effect, which takes into account the repeated observations within each patient.

Estimation of Integral Curves from Noisy Diffusion Tensor Data

✦Lyudmila Sakhanenko, Michigan State University, Wells Hall, Dept. of Statistics and Probability, East Lansing, MI 48824, *luda@stt.msu.edu*

Key Words: diffusion tensor imaging, kernel smoothing, asymptotic normality, integral curve, limit theorems

Consider a tensor field in a bounded subset in R^d. Its components are slopes in a certain system of regression equations. The problem is to estimate the integral curve starting at a given point and driven by the principle eigenvector field corresponding to this tensor field. We develop a three-step estimation procedure based on a kernel estimator for the tensor field, followed by a plug-in estimator for the vector field and a plug-in estimator for the integral curve. Our work utilizes ideas of Koltchinskii, Sakhanenko and Cai (2007) in a more complicated, more general and more realistic than theirs model. The resulting estimator for the integral curve is asymptotically normal. We derive the differential and integral equations for the mean and the covariance of the limiting Gaussian process. This provides a method for tracking the integral curve together with its mean and covariance and allows to construct confidence ellipsoids for fixed points on the curve. This problem is motivated by Diffusion Tensor Magnetic Resonance Imaging (DT-MRI), a popular brain imaging technique. We provide a statistical framework for assessing uncertainty in images based on diffusion tensor data.

A Lifting Scheme For Multiscale Kernel Smoothing

✦Maarten Jansen, UniversitÈ Libre de Bruxelles, Boulevard du Triomphe, Campus Plaine, CP213, Brussels, B-1050 Belgium, *maarten.jansen@ulb.ac.be*

Key Words: lifting, kernel, wavelet, Laplacian pyramid, smoothing

The lifting scheme provides a framework for the design of wavelet transforms for observations on inequisdant covariates. With a design of the transform based on interpolation, the irregularity may invoke numerical problems, when adjacent observations are much closer to each other than to other neighbours. As an alternative, we propose to construct a transform based on smoothing techniques, kernel smoothing being an example. Smoothing is far less sensitive to the errors and the geometry of irregular observations. Unfortunately, smoothing is not an interpolating operation, thereby leading to discontinuous, unsmooth basis functions. The contradiction in the objectives of smoothness and numerical condition can only be solved by the introduction of a slight redundancy in the transform. This leads to a combination of the lifting scheme and Burt-Adelson's Laplacian pyramid. We discuss the design of such a transform, and more specifically the primal moment condition (stating that wavelets should have zero integrals for

stability in ell-2), the dual moment condition (stating that constant inputs should have zero detail coefficients, for the sake of sparsity) and perfect reconstruction.

Minimax-Optimal Rates For Sparse Additive Models Over Kernel Classes Via Convex Programming

✦Garvesh Raskutti, University of California, Berkeley, , *garveshr@stat.berkeley.edu*

Key Words: high-dimensional, sparse additive models, minimax, convex optimization

Sparse additive models are families of d-variate functions that have the additive decomposition \mbox{$f^* = \sum_{j \in S} f^*_j$,} where S is a unknown subset of cardinality $s \ll d$. We consider the case where each component function f^*_j lies in a reproducing kernel Hilbert space, and analyze a simple kernel-based convex program for estimating the unknown function f^*. Working within a high-dimensional framework that allows both the dimension d and sparsity s to scale, we derive convergence rates in the $L^2(\mathbb{P})$ and $L^2(\mathbb{P}_n)$ norms. These rates consist of two terms: a \emph{subset selection term} of the order $\frac{s \log d}{n}$, corresponding to the difficulty of finding the unknown s-sized subset, and an \emph{estimation error} term of the order $s \, \nu_n^2$, where ν_n^2 is the optimal rate for estimating an univariate function within the RKHS. We complement these achievable results by deriving minimax lower bounds on the $L^2(\mathbb{P})$ error, thereby showing that our method is optimal up to constant factors for sub-linear sparsity $s = o(d)$. Thus, we obtain optimal minimax rates for classes of sparse additive models.

Feature Based Image Registration Using Non-Degenerate Pixels

✦Chen Xing, University of Minnesota -- Twin Cities, 471 Ford Hall, 224 Church Street S.E., Minneapolis, MN 55455, *xingx011@umn.edu*; Peihua Qiu, University of Minnesota -- Twin Cities

Key Words: image registration, degenerate, local smoothing, image matching, feature extraction, thin plate splines

Image registration (IR) aims to geometrically match one image to another. It is extensively used in many imaging applications. Among many existing IR methods, one widely used group of methods are feature-based. By a feature-based method, a number of relevant image features are first extracted from the two images, respectively, and then a geometric matching transformation is found to best match the two sets of features. However, proper identification and extraction of image features turns out to be a challenging task. Generally speaking, a good image feature extraction method should have the following two properties: (i) the identified image features should provide us enough information to approximate the geometric matching transformation accurately, and (ii) they should be easy to identify by a computer algorithm so that the entire feature extraction procedure is computer automatic. In this paper, a new type of image features is studied, which has the two properties described above. Together with the widely used thin plate spline (TPS) geometric transformation model, it is shown in the paper that our feature-based IR method works effectively in various cases.

513 Advances in Design

Section on Statistical Computing, International Indian Statistical Association, Section on Quality and Productivity, Section for Statistical Programmers and Analysts

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Optimal Designs for Rational Function Regression

✦David Papp, Rutgers University, Piscataway, NJ 08854, *dpapp@rutcor.rutgers.edu*

Key Words: Optimal design, Rational function regression, Semidefinite programming, Nonlinear regression

We consider optimal non-sequential designs for a large class of (linear and nonlinear) regression models involving polynomials and rational functions with heteroscedastic noise also given by a polynomial or rational weight function. The proposed method generates a polynomial whose zeros are the support points of the optimal approximate design, and generalizes a number of previously known results of the same flavor. The method is based on a mathematical optimization model that can incorporate various criteria of optimality and can be solved very efficiently by well established numerical optimization methods. In contrast to previous optimization-based methods proposed for similar design problems, it also has theoretical guarantee of its efficiency. After discussing linear models, applications for finding locally optimal designs for nonlinear regression models involving rational functions are presented, then extensions to robust regression designs, and trigonometric regression are shown. As a corollary, an upper bound on the size of the support set of the minimally-supported optimal designs is also found.

The Balanced Incomplete Two-Way Block Design

✦M. SHAMSUDDIN, KING ABDULAZIZ UNIVERSITY, PROFESSOR, DEPARTMENT OF STATISTICS, JEDDAH, International 80203 SAUDI ARABIA, *dr_md_shams2004@yahoo.com*; Mian Arif Shams Adnan, JAHANGIRNAGAR UNIVERSITY

Key Words: Balanced Incomplete Block Design, Factorial Design, Row-Column Design

The Balanced Incomplete Row-Column Design originating from partially confounded orthogonally replicated row-column factorial designs is called the Balanced Incomplete Two-way Block Design. Here the rows and columns are contemporarily incorporated for eliminating two-way heterogeneity. Moreover, the adjusted treatment sum of squares of this design is same to the total of the sum of squares of all factorial effects. Some examples are cited to illustrate the techniques of the entire methodology.

Emulating A Computationally Expensive Likelihood Function: A Sequential Design Approach

✦Antony Overstall, Southampton Statistical Sciences Research Institute (S3RI), University of Southampton, Highfield, Southampton, International SO17 2BJ United Kingdom, *amo105@soton.ac.uk*; David Woods, Southampton Statistical Sciences

Research Institute (S3RI)

Key Words: computationally expensive likelihood, computer experiments, sequential design, emulator

The likelihood function plays a crucial role in statistics; both in the classical and Bayesian paradigms. In some cases, the likelihood function can be computationally expensive, making statistical inference time consuming or even impossible. Our strategy for this problem is to construct a computationally cheap approximation to the likelihood function and use this emulator as a surrogate for the true likelihood in statistical inference. This is an application of computer experiment methodology - a current area of active research. The emulator is constructed on the basis of a set of evaluations of the likelihood at different points in the parameter space. To select these points we use an iterative sequential design approach which uses the emulator at the current iteration to choose where next to evaluate the likelihood.

Some Iterative Procedures For Constructing Optimal Designs

✦Saumen Mandal, University of Manitoba, Department of Statistics, 338 Machray Hall, Winnipeg, MB R3T 2N2 Canada, *saumen_mandal@umanitoba.ca*

Key Words: Algorithms, Directional derivatives, Optimality conditions, Optimizing distributions, Optimal designs

We consider the problem of computing designs which optimize some design criteria subject to the basic constraints on the design weights. Examples of such problems include parameter estimation, adaptive design and stratified sampling. We search for an optimal distribution in a design space which is a discretization of a continuous space. In order to construct the optimal designs we use a class of algorithms, indexed by a function which depends on the derivatives of the criterion function and satisfies some conditions. The function may depend on one or more free positive parameters. These algorithms possess some suitable properties and are shown to satisfy the basic constraints on the design weights of non-negativity and summation to unity. We first investigate the performance of the algorithm by constructing D-optimal designs. We then explore the properties of the partial and directional derivatives of the criterion function and develop strategies for improving convergence rates of the algorithms.

Avoiding The Notorious Ill-Conditioning Of The Correlation Matrix Used In Gaussian-Processes Design Of Experiments By Use Of Symbolic Sensitivity Analysis

✦Selden Crary, Crary Group, Inc., 535 Everett Ave. Apt. 410, Palo Alto, CA 94301, *selden_crary@yahoo.com*; Rachel Johnson, Naval Postgraduate School

Key Words: design of computer experiments, correlation matrix, gaussian processes, ill-conditioning, symbolic sensitivity analysis, twin points

We use a symbolic sensitivity analysis to circumvent the notorious ill-conditioning of the correlation matrix V used for computing the integrated mean squared error (IMSE) in gaussian-processes design of experiments. The steps are as follows: (1) expand V in powers of the squared, weighted separation Delta between the pair of proximal design points x1 and x2; (2) isolate the occurrences of Delta to a 2x2 block and rename the matrix Vb; (3) invert Vb in hybrid-numeric-symbolic form; (4) expand the inverse in powers of Delta; (5) compute IMSE as a power law in Delta; and (6) take the limit Delta goes to zero. The potentially problematic terms of the inverse, which include terms proportional to 1/Delta, cancel identically. This method allows the computation of IMSE for designs with a pair of proximal points, without resorting to extended-precision arithmetic or other severe measures. The method was applied to a putatively twin-point design described in the engineering literature. That twin-point design was found, in the present study, to occupy a minimum in the design space, thus lending support to its identification as the optimal design for the problem.

Design Of A Functional Programming For Statistical Analysis

✦Babubhai V. Shah, SAFAL Institute Inc., 22 Autumn Woods Dr., Durham, NC 27713, *babushah@earthlink.net*

Key Words: Functional Programming, statistical analysis, multi-processor system, parallel computation

Object oriented languages (APL, S-plus, R) treat multidimensional array as an object, and have bee useful in statistical analysis of data. However, these languages retain Von Neumann style semantics. For implementing a functional programming language an abstract concept of "virtual" dimension is introduced. Its semantics are connected to mathematical sequences. Standard statistical data sets used by statistical procedures in SAS, SPSS, SUDAAN is a special case. We define functions for create these objects and define other objects. A few new primitive functions and a generalization of known scalar and matrix functions make it possible to define most statistical analysis. We illustrate how these concepts help define simulations, iterations as well as summarizations. These definitions expose the potential parallel computation to enable implementation on multi-processor systems.

Optimal Designs For Factorial Experiments With Binary Response

✦Abhyuday Mandal, University of Georgia, , *amandal@stat.uga.edu*; Jie Yang, University of Illinois at Chicago; Dibyen Majumdar, University of Illinois at Chicago

Key Words: generalized linear model, factorial design, D-optimality, information matrix, Bayesian designs, exchange algorithms

We consider the problem of obtaining locally D-optimal designs for factorial experiments with qualitative factors with binary response and identify optimal allocation of runs for 2^k and 2^{k-p} designs. We develop efficient numerical techniques for solving this very high dimensional optimization problem. We also identify the Bayesian D-optimal designs and compare their performances with locally D-optimal designs.

514 Statistical Issues and Solutions

General Methodology, Biometrics Section, Biopharmaceutical Section, Business and Economic Statistics Section, Section on Statistics and the Environment, Section on Statistics in Epidemiology

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Inner Envelopes: Efficient Estimation In Multivariate Linear Regression

✦Zhihua Su, University of Minnesota, 472 Ford Hall, 224 Church St. SE, Minneapolis, MN 55455 United States, *suzhihua@stat.umn.edu*; R. Dennis Cook, University of Minnesota

Key Words: Dimension reduction, Envelope model, Reducing subspaces, Grassmann manifolds

In this paper, we propose a new class of models, called the inner envelope model which leads to efficient estimation in the context of multivariate linear regression. The asymptotic distribution and the consistency of its maximum likelihood estimators are established. Simulation studies and examples both show that the efficiency gains can be massive. Compared to the envelope model introduced by Cook et al. (2010), the inner envelope model has a different starting point on the mechanism of improving the efficiency, and it can produce substantial gains in the situations where the envelope model offers no gains. The construction of the inner envelope model also opens a new frontier on the way in which reducing subspaces can be used to improve efficiency in multivariate problems.

Some Thoughts In Multiple Hypothesis Testings

✦Yi Cao, University of Texas at San Antonio, 6061 DE ZAVALA RD APT 1521, San Antonio, TX 78249 U.S., *yi.cao@utsa.edu*; Keying Ye, University of Texas at San Antonio

Key Words: multiple testing, k family-wise error rate, false discovery rate

Two classical metrics in analyzing multiple hypothesis testings are family- wise error rate (FWE) and all-pair power which are widely used in many years until false discovery rate (FDR) and average power catch our attention in error measurement and power enhancement. In this research, the k-FWER and k-pair power, two less conservative notions of error rate and power are demonstrated to be tightly connected with FDR and the average power. Based on this connection, various current hypothesis testing procedures such as Benjamin Hochberg, Storey, Hommel,Holm's procedure and so on are compared by a series of simulations. We also notice that as the number of hypotheses increases, k needs to be adjusted in a more flexible way rather than a preset value. A new beta-FWER in which beta is made as a function of the number of hypotheses is proposed to balance the error rate and power. A stepwise beta-FWER controlling procedure is also constructed. Intensive empirical simulations show that this procedure is more reasonable in general condition and advantageous under some special cases than current k-FWER controlling procedures and is no worse than FDR controlling procedures in terms of power

Likelihood-Based Instrument Variable Analysis For Comparing Survival Of Prostate Cancer Patients Receiving Hormone Therapy Vs. Conservative Management

✦Anjun Cao, University of Medicine and Dentistry of New Jersey, 155 Matthews Farm Road, Belle Mead, NJ 08502, *ajco@novonordisk.com*; Dirk F Moore, University of Medicine and Dentistry of New Jersey

Key Words: IVA, covariate adjustment, prostate-specific survival, PSA level

Instrument variable analysis (IVA) is a method for comparing interventions using observational data. Unlike classical methods of covariate adjustment, it can adjust for both measured and unmeasured covariates, provided that essential assumptions are satisfied. The most important assumption is that an instrument can be found that is associated with the outcome of interest only through its association with the intervention. We present a two-stage likelihood-based IVA procedure for comparing survival data between two interventions. We show that the method can produce unbiased estimates of survival differences if the distribution of an unobserved confounder is known. We use the method to compare prostate-specific survival of patients receiving hormone therapy vs. conservative management, where the unobserved confounder is PSA level.

Sample Size Guidelines When Using A Validation Study To Adjust For Measurement Error Bias In Linear Regression

✦John W. Rogers, Westat, 1600 Research Blvd., Rockville, MD 20850, *johnrogers@westat.com*

Key Words: Regression calibration, Multiple imputation

From a set of N observations of Y and X, regression parameters predicting Y will be biased when X is measured with error. Unbiased slope estimates can be obtained if an estimated measurement error variance is available from a validation sample with paired measurements X and T, the "gold standard" for the quantity measured by X. Such situations often arise in environmental sampling, such as using a sampling badge as a surrogate for personal exposure monitoring or a pesticide use questionnaire as a surrogate for pesticide exposure. An internal validation sample may be collected as part of the N observations or an external estimate can be obtained from a separate study. The uncertainty in the measurement error variance contributes uncertainty to the bias corrected slopes. That uncertainty needs to be included when determining sample sizes. This paper presents sample size guidelines for the size of the validation sample relative to the entire sample. The suggested sample sizes depend on the cor(X,Y), cor(Y,T), and the sample size that would be used if the gold standard was used for all cases. Guidelines are presented for both internal and external validation samples.

Posterior Predictive Distributions In Meta-Analysis: An Updated Hunter-Schmidt Analysis Using Bayesian Hierarchical Models

✦David G Whiting, Brigham Young University, 223C TMCB, Brigham Young University, Provo, UT 84602, *whiting@byu.edu*; Michael D Ulrich, The RBL Group

Key Words: posterior predictive distributions, meta-analysis, correlations

The traditional Hunter-Schmidt approach to meta-analysis includes fixed effect, random effect, and empirical Bayesian models for combining correlations from multiple studies. When modeling correlations, each of these approaches takes into account corrections for reliability. However, they do not fully account for the uncertainty introduced by estimating the reliability coefficients themselves. We propose a Bayesian hierarchical model for correlation meta-analysis which fully accounts for these sources of model variability. We demonstrate how the posterior predictive distributions of the study correlations yields more information for drawing overall conclusions. We examine a highly-cited meta-analysis from the organizational behavior literature, showing how its conclusions would have been different had reliability uncertainty been accounted for.

Applications Of Matrix Approximation By Polynomial Factor Analysis

✦Kekona Kai Sorenson, UCLA Department of Statistics, 8125 Math Sciences Bldg., Box 951554, Los Angeles, CA 90095-1554, *kekona@gmail.com*

Key Words: Nonlinear Factor Analysis, Alternating Least Squares Fitting, Jacobi-Givens Plane Rotations

Factor analysis is a valuable exploratory multivariate data analysis technique that has blossomed over the past 100 years into the preferred technique for a wide range of disciplines. Even though factor analysis has been well received in the applied sciences, statisticians have been hesitant to adopt the technique because, among other issues, it is has traditionally been limited to a linear model. Much work has been done to date to augment the proliferation of nonlinear extensions to the standard factor analysis model, and in this paper we will continue on that path. Inspired by the work done in Gi? and by Jan de Leeuw, we propose a technique to ?t multivariate polynomials of a number of orthogonal factors to a data matrix. The algorithm is of the Alternating Least Squares type, in which minimization over the loadings for a given set of factors and minimization over a set of factor scores for a given set of loadings are alternated until a convergence criteria is satis?ed. For this presentation I will focus on discussing the speci?c application to environmental data sets, as well as a quick discussion of applications to psychology and education.

515 Advanced Topics in Estimators and Confidence Intervals

Section on Physical and Engineering Sciences

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Point Estimation Of The Central Direction Of Random Rotations In So(3)

✦Bryan Stanfill, Iowa State University, IA , *stanfill@iastate.edu*; Ulrike Genschel, Iowa State University; Heike Hofmann, Iowa State University; Dan Nordman, Iowa State University

Key Words: Random rotations, Central direction, Uniform-axis random spin distributions, Distance metrics

Random rotations in SO(3) (the space of 3x3 orthogonal matrices with determinant 1) are used to model three-dimensional orientations in fields such as computer science, kinesiology, material science, and even astrophysics. Currently, several estimators of the central direction are suggested in the literature but little attention has been paid to a comparison study examining which provides the best estimate. We compare four favorable estimators across three commonly used uniform-axis random spin distributions. Three of the estimators have been studied in the literature thus far at varying depths whereas the fourth estimator is adapted to rotation matrices for the first time. For the von Mises distribution the normalized spatial median yields preferable results whereas for the Cayley and Fisher distributions an estimator based on the L2 norm is preferred. These findings, when applied to published papers in the material sciences, prove there is room for improvement in the estimation process used presently.

Stein'S Initial Estimator Of A Multivariate Normal Mean

✦Jerome P. Keating, University of Texas at San Antonio, Department of Management Science & Statistics, San Antonio, TX 78249-0704, *jerome.keating@utsa.edu*; Robert L. Mason, Southwest Research Institute

Key Words: Stein rule, Hyperbolic geometry, Astrophysics

In 1956, Charles Stein introduced the simple version of his famous Stein rule estimator, which has now been generalized in numerous ways. However, from the text of the original paper it is clear that Professor Stein first considered an alternative version, based on the positive-part of the square-root of the coefficient of his famous discovery. We investigate his initial-rule and its connection with hyperbolic geometry. The initial Stein-rule is shown to be a contraction toward the origin predicated on distance in hyperbolic space. The standard estimator of a multivariate normal mean is shown to be inadmissible to Stein's initial-rule as well. Stein's initial-rule has practical application in situations where the Stein rule appears to shrink the components of the estimated normal mean too much toward the origin (or any other specified point). Likewise, given its roots in hyperbolic geometry, Stein's initial-rule should have many applications in astrophysics as well.

The Liu, Ridge, and Least Square Estimators: A Comparison

✦Marvin Haskel Jack Gruber, Rochester Institute of Technology, 130 Rivermeadow Drive, Rochester, NY 14623, *mjgsma@rit.edu*

Key Words: Zellner's balanced loss function, mean square error, shrinkage estimator, optimal estimator

In a 1993 paper Liu proposed an estimator that is similar in form but different from the ridge regression estimator of Hoerl and Kennard. More recently in 2007 Ozkale and Kaciranlar proposed a two parameter variation of the Liu estimator. We will derive a generalization of the two parameter Liu estimator that is optimal with respect to Zellner's balanced loss function (ZBLF) for prior assumptions about the first and second moment. This Liu type estimator specializes to the generalized ridge regression estimator for the classical loss function(CLF). The Bayes and the frequentist risks with respect to both ZBLF and CLF will be compared for the Liu type estimator, the ridge estimator and the least square estimator.

Pairwise Maximum Likelihood Procedure For Estimating Shift In Two-Sample Location Problem

✦Feridun Tasdan, Western Illinois University, Department of Mathematics, 1 University Circle, Macomb, IL 61455 USA, *f-tasdan@wiu.edu*

Key Words: Two-Sample Problem, Pairwise Maximum Likelihood, Likelihood Ratio, Shift Parameter

This study is about maximum likelihood estimation of shift parameter using pairwise differences in two-sample location problem. The proposed procedure uses pairwise differences of data to find a distribution function of the differences. Hence, using the distribution function of the pairwise differences, one can find a likelihood function with respect to the shift parameter. By using the likelihood function, a MLE of shift parameter will be found. Moreover, it will be shown that the proposed procedure is equivalent to the regular likelihood ratio test. An asymptotic level hypothesis test and confidence interval will be investigated for the proposed estimator. The study ends with a bootstrap simulation study which is performed to show the efficiency results of the proposed estimator.

Interchangeability Among Confidence, Prediction And Tolerance Intervals

✦Harry Yang, MedImmune, One MedImmune Way, Gaithersburg, MD 20878, *YangH@medimmune.com*; Lanju Zhang, MedImmune

Key Words: Confidence interval, Predication interval, Shelf life, Tolerance interval

Concepts of confidence, prediction and tolerance intervals are discussed. Assuming that the sampling distribution is normal, it is shown that the confidence interval (CI) and prediction interval (PI) are special cases of the tolerance interval (TI). Under such unified perspective, these interval approaches can be used and interpreted interchangeably. Two exact methods, one based on generalized confidence interval and the other exact integration, for determining tolerance bounds are developed. A simulation study is conducted to compare coverage probability of the methods to those currently in use. The unified view also makes it possible to evaluate different interval approaches in the same statistical context. An application of the unified method to drug shelf life estimation is provided.

Shortest Expected Length Confidence Interval For The Power Of The T-Test

Subhabrata Chakraborti, The University of Alabama; Greg Michaelson, The University of Alabama.; ✦Amanda McCracken, The University of Alabama, , *amanda.mccracken@gmail.com*

Key Words: power, confidence interval, shortest length, effect size

The t-test is among the most popular statistical tests, and its power is of concern in numerous applications. However, the power of a t-test depends on the unknown population standard deviation, s. This problem is often dealt with by substituting the sample standard deviation s for s, but doing so results in an estimated power which is a random variable. Since estimating power can have important practical implications, it is of interest to calculate a confidence interval on power. Tarasinka [2005] suggested a method of calculating such a confidence interval which results in the shortest length interval for power for an observed value of s. However, since s is a random variable, the”optimal” endpoints vary with s, and the resulting intervals have lower than nominal coverage, as demonstrated by Gilliland and Li [2008]. We propose a method of obtaining the shortest expected length confidence interval for power by accounting for the variation in s. The proposed interval maintains nominal coverage and is robust over various estimates of the standard deviation. Comparisons are made with an existing proposal by Tarasinka [2005].

Conditional Confidence Intervals Of The Normal Variance And Its Application To The Process Capability Index Cp

✦Chien-Pai Han, University of Texas at Arlington, Department of Mathematics, Arlington, TX 76019, *cphan@uta.edu*; Jianchun Zhang, University of Texas at Arlington

Key Words: Conditional Confidence Interval, Preliminary Test, Process Capabality Index, Coverage Probability

To construct confidence intervals for the normal variance is an important inference procedure. When the experimenter has prior information about the variance, it is advantageous to use the prior information. When the experimenter is uncertain about the prior information, a preliminary test can be used to verify the validity of the information. We consider the construction of a conditional confidence interval for the variance after the rejection of the null hypothesis that the population variance is equal to a given prior value. Comparisons of the conditional confidence intervals and unconditional confidence intervals are made and the conditional coverage probabilities of the unconditional confidence intervals are investigated. Since the process capability index Cp involves only one parameter that is the process standard deviation, the conditional confidence interval for Cp can also similarly obtained.

516 Optimizing Sample Designs

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Designing Minimum-Cost Multi-Stage Sample Designs

✦Matthias Ganninger, GESIS - Leibniz-Insitute for the Social Sciences, P.O. Box 122155, Mannheim, International 68072 GERMANY, *matthias.ganninger@gesis.org*

Key Words: design effects, multi-stage design, costs, minimum

In cross-national sample surveys like the ESS, a huge variety of sample designs is often applied in participating countries. In order to achieve estimates of comparable precision, the samples drawn according to these different sampling schemes must be of equivalent effective sample sizes, n_eff = n/deff, where n is the net sample size and deff is the design effect. As deff, among another parameter, depends on the average cluster size b, increasing the number of sampled clusters, ceteris paribus, decreases the design effect and hence increases n_eff. The presentation will show that, at a given linear cost structure (costs per interview and costs per sampled cluster), there exists an optimal number of clusters to sample so that a pre-defined effective sample size is exactly achieved - at minimum total costs.

A Generalized Epsem Two-Phase Design For Domain Estimation

Avinash C Singh, NORC at the University of Chicago; ✦Rachel Harter, NORC at the University of Chicago, 55 E. Monroe St., Chicago, IL 60603, *harter-rachel@norc.org*

Key Words: two-phase, self-weighting, EPSEM, domain estimation

Composite measures of size can be used to select primary sampling units in a two-phase design such that multiple subdomains are self-weighting. This method can be generalized to situations where the PSU probabilities are prescribed and the sample allocations to domains within PSUs are adjusted to achieve self-weighting domains. The method is illustrated for an area probability sample of housing units.

Assigning Psus To A Stratification Psu

✦Susan King, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Room 3650, Washington, DC 20212, *king.susan@bls.gov*; John John Schilp, U.S. Bureau of Labor Statistics; Erik Bergmann, U.S. Bureau of Labor Statistics

Key Words: clustering, integer programming, assignment algorithm, stratification PSU, Consumer Expenditure Survey, Consumer Price Index

After every decennial census, many surveys including the Consumer Price Index (CPI) and the Consumer Expenditure Survey (CE) redefine their primary sampling units (PSUs), which are sets of contiguous counties. Since the CE survey is used to weight the CPI, the two surveys use a common set of PSUs. There are two types of PSUs: self representing and non-self representing PSUs. Self representing PSUs are selected with certainty, whereas non-self representing PSUs are grouped into a stratification PSU and one PSU is randomly selected to represent the stratification PSU. To minimize survey variance, the stratification PSUs should be homogeneous and have approximately equal populations. This is a constrained clustering problem and is solved using heuristic algorithms. This paper presents a new heuristic solution procedure that uses a "pseudo" assignment algorithm to assign PSUs to a stratification PSU. This heuristic procedure found a lower within cluster variability, Trace (W), than other procedures.

An Enhanced Approach For Solving The Overlap Problem In Dual Frame Rdd Surveys

✦Sohyung Park, Dongguk University, 26,3 Pil-dong, Jung-gu, Seoul, 100-715 South Korea, *12astro@naver.com*; Sun-Woong Kim, Dongguk University; Michael Traugott, University of Michigan; Sang-Kyung Lee , Hyundae Research Institute

Key Words: household number, shared number, personal number, person-level weighting

Dual frame RDD designs of landline and cell phone numbers to eliminate the coverage bias due to cell-only populations have been popular in many countries, but they are subject to the overlap and overrepresentation problem. One of the solutions to the problem is to use a compensatory weight for unequal probabilities of selection, but weighting is often inappropriate because of lack of information on actual phone usage. We consider a two-stage procedure, illustrated by Lepkowski and Kim (2005), different from a conventional within-household selection. First, with the assumption that both landline and cell numbers are 1) for the household, 2) shared, or 3) personal numbers, a few questions are used to identify such status from the informant for each number selected from the frame. Second, if the phone number is for the household or a shared number, one eligible person using the number is randomly chosen and asked to provide the information on other phone numbers that could have been reached to the person. This approach was successfully applied to a national survey in Korea. Several person-level weighting strategies based on phone ownership data from the survey are examined.

Updating The Measures Of Size Of Local Areas Late In The Decade Using Usps Address Lists

✦Sylvia Dohrmann, Westat, 1600 Research Blvd, Rockville, MD 20850, *sylviadohrmann@westat.com*; Lin Li, Westat; Leyla K. Mohadjer, Westat

Key Words: USPS, area sample, measure of size (MOS)

Area household surveys conducted in the United States most often rely on data from the Census Bureau to calculate measures of size (MOS) of secondary sampling units (SSUs or segments). Yet, late in the decade demographic or housing data from the last decennial census are likely to be inaccurate in local areas with considerable growth or demographic shifts since the census taking and intercensal estimate are not available at the required level. Address lists available from the United States Postal Service (USPS) have been incorporated into survey sample designs in various capacities over the past decade: telephone surveys are using these lists as a first phase of selection; in-person area surveys are using them in place of the traditional address listing process; and mail surveys are becoming much more prevalent. In all these instances, the lists are being used as sampling frames. This paper will present two implementations of the use of addresses lists as a means of updating segment MOS late in the decade and demonstrate their effectiveness.

Calibrating Non-Probability Online Samples With Probability Samples Using Early Adopter Characteristics

✦Charles DiSogra, Knowledge Networks, Inc, 1350 Willow Road, Suite 102, Menlo Park, CA 94025, *cdisogra@knowledgenetworks.com*; J. Michael Dennis, Knowledge Networks, Inc.; Elisa Chan, Knowledge Networks, Inc.

Key Words: Online surveys, Web surveys, probability panels, calibration, weighting

Representative study samples drawn from the Knowledge Networks (KN) probability-based web panel, after post-stratification weighting, reliably generalize to the population of interest. Due to finite panel size, there are instances of too few panel members to meet sample size requirements. In such cases, a carefully designed supplemental sample from an opt-in panel source may be added. When both samples are profiled with 5 questions on early adopter (EA) behavior, non-probability opt-in samples more often tend to be strongly EAs compared to probability samples. When a weighted KN panel sample is blended with opt-in sample and re-weighted, the EA profile of the resulting blended sample can skew toward the EA profile of the opt-in cases.

This bias will be demonstrated with results from several blended survey samples. A similar bias may also be seen to affect key study estimates. A calibration approach incorporates one or more EA questions as weighting variables in a raking procedure and thus, this measured bias may be largely minimized, if not eliminated. Results from several study samples will be shown to successfully demonstrate this approach.

Non-Response Follow-Up Allocation For Domains

✦Harold Mantel, Statistics Canada, 16-H R.H. Coats Building, Tunney's Pasture, Ottawa, ON K1A 0T6 Canada, *harold.mantel@statcan.gc.ca*; Mike Hidiroglou, Statistics Canada

Key Words: non-response, follow-up, allocation, domains

In this paper we consider the problem how to optimally allocate a subsample of non-respondents from an initial sample for follow-up, when we want to estimate totals or means for domains. We consider different scenarios for what is known. At one extreme, we may know the domain of every unit in the initial sample, whether it responds or not. Alternatively, we may only know the domain of the respondent units, but have some auxiliary information such as the domain sizes. We may also consider different scenarios for knowledge of how the domains interact with strata. We assume that follow-up is intensive, so that all of the non-respondents in the follow-up sub-sample will become respondents, although this is unrealistic in practice. We then derive expressions for the variance. The optimal allocation is then the one that minimizes some function of the domain variances, such as maximum or total. After deriving an optimal allocation and its associated variance, we can go further and try to find an optimal balance between allocation to the initial sample and the follow-up, under an assumed cost model.

517 Statistical Models for Genetic Analysis

SSC, Section on Nonparametric Statistics, ENAR, Section on Statistics in Epidemiology, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

A Generalized Codon-Based Model Of Nucleotide Substitution For Protein-Coding Dna Sequences

✦Toby Kenney, Dalhousie University, Dept. of Math. & Stat.,, Dalhousie University, Halifax, NS B3H 3J5 Canada, *tkenney@mathstat.dal.ca*; Hong Gu, Dalhousie University

Key Words: Protein-coding DNA Sequences, codon-based model, phylogenetic inference, molecular evolution, general time-reversible model, Newton-Raphson algorithm

A generalized codon-based model for protein-coding DNA sequences is proposed for phylogenetic analysis. This model framework provides a unified framework for existing codon models (and also for DNA and amino acid models). Furthermore, it offers greater flexibility in the choice of rate matrix, allowing existing models to be easily extended to incorporate more of the possible driving forces in molecular evolution, such as structure information and amino acid properties. Our generalized codon model allows the rate matrix to be either sparse, as in the current codon models or patterned to allow all sorts of multiple changes, or anywhere in between. We provide a software package called Codon Optimal Likelihood Discoverer (COLD) to implement these proposed generalized codon models. COLD uses the standard Newton-Raphson algorithm to optimize the likelihood to get the estimation of all the parameters. We demonstrate how our model framework allows the parametrisation to be determined by statistical model selection, and statistical inference can be performed for the parameters involved in the model based on standard likelihood theory.

Identifying Differential Gene Sets Using The Linear Combination Of Genes With Maximum Auc

✦Chen-An Tsai, Graduate Institute of Biostatistics & Biostatistics Center. China Medical University, Taichung, International Taiwan, *catsai@mail.cmu.edu.tw*; Zhanfeng Wang, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan; Yuan-chin Ivan Chang, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Key Words: Gene Set Enrichment Analysis, discriminatory power, ROC, AUC

Gene set enrichment analyses (GSEA) provide a useful and powerful approach to identify differentially expressed gene sets with prior biological knowledge. Several GSEA algorithms have been proposed to perform enrichment analyses on groups of genes. However, little attention has been given to the discriminatory power of gene sets. Thus, it is of great interest to identify which differential gene sets are strongly associated with phenotypic class distinction ability by integrating gene expression data with prior biological knowledge. We propose two methods to identify differential gene sets using the area under the receiver operating characteristic (ROC) curve (AUC) of linear risk scores of gene sets, which are obtained through a parsimonious threshold-independent gene selection method within gene sets. The p-values of AUC-based statistics and the AUC values obtained from cross-validation of the linear risk scores are calculated, and used as indexes to identify differential gene sets. The discrimination powers of gene sets are summarized and gene sets that possess discrimination power are selected via a prescribed p-value threshold or a predefined cross-validation AUC threshold.

A Hierarchical Dirichlet Process Mixture Model For Genotype Calling: Combining Snp Calling Algorithm (Within Sample) With Sample Genotyping Strategy (Each Snp)

✦Gengxin Li, Yale University, STE 503, 300 George ST, New Haven, CT 06511-6660 USA, *gengxin.li@yale.edu*; Hongyu Zhao, Yale University

Key Words: Dirichlet Process, Genotype Calling, mixture model

Large scale genotyping is a crucial process to assign genotype to hybridization data through calling algorithms. With the development of the illumine chip technique, a number of genotyping methods have been developed recently. The classical genotyping calls for a SNP are made by GenCall with GenTrain clustering algorithm that requires a large number of samples to make sure the calling accuracy. When rare SNPs are commonly observed in the large scale genotyping, it is extremely difficult to perform the genotype calling. GenoSNP based on within-sample SNP genotyping was explored to address this challenge. Because of the richness of DNA information within sample, its calling result greatly reduces the missing rate, especially for some rare SNPs and samples. However, GenoSNP fails the Hardy-Weinberg equilibrium tests for a great many SNPs due to the assumption that SNPs within-sample is hard to satisfy. In this paper, we propose an efficient Hierarchical Dirichlet Process Mixture Model to combine the superiorities of both calling methods. While comparing with GenCall and GenoSNP, the better performance of our discussed method is demonstrated in the simulation study and real case.

Exact Tests Of Hardy-Weinberg Equilibrium In Case-Parent Triad Design And Algebraic Statistics

✦Subramanyam Kasala, University of North Carolina at Wilmington, Department of Mathematics and Statistics, Wilmington, NC 28403, *kasalas@uncw.edu*; Marepalli B Rao, University of Cincinnati; Marepalli B Rao, University of Cincinnati

Key Words: Triad Design, Alleles, Hardy-Weinberg Equilibrium, p value, Markov Base, MCMC

Case-Parent Triad Designs are useful in genetic studies. Typically, in a triad desgin, association between a bi-allelic marker and a complex disease will be under focus. A triad consists of a child and child's parents. Triad data consist of genotypes of the members of the triads in the sample. The focus of the presentation is developing an exact test of Hardy-Weinberg Equilibrium (HWE) at the marker a la Fisher. In the literature, tests of HWE are conducted obliquely. Here, we postulate the null hypothesis of HWE directly and the hypothesis is composiste with one nuisance parameter. Our exact test involves algebraic statistics, Markov bases, and MCMC. The key idea is to look at the fiber of all possible triad data sets with allele frequencies equal to those of the given triad data set. It turns out that the fiber is a certain collection of 2x15 matrices with non-negative integer entries and prescribed row sums. The size of a fiber is usually very large. We will build a Markov base, which is small, for the fiber. The given triad data set and the Markov base will generate the entire fiber. We will then develop an MCMC algorithm to estimate the p value under the equilibrium.

Gene Enrichment In Genome-Wide Association Studies

✦Stephen Aaron Stanhope, University of Chicago, 1813 W. Cortland St., Apt. 2F, Chicago, IL 60622, *sstanhop@bsd.uchicago.edu*; Mark Abney, University of Chicago

Key Words: Genome-wide association study, Gene enrichment

In a genome-wide association study (GWAS), hundreds of thousands of markers are genotyped across hundreds of cases and controls, and each marker is tested for disease associations. After controlling for multiple comparisons, it can be difficult to detect loci with small effects on disease risk. It is therefore attractive to limit the number of tests by analyzing groups of related markers. We describe an enrichment study of hypertension in the Hutterites, a population related through a 13-generation pedigree. We aggregate score statistics from a GWAS into gene scores using the array annotation and a scoring function. Because of linkage disequilibrium, under the null of no markers within a gene being associated with the disease, the distribution of the gene score is that of a function of correlated chi-squared distributed random variables. We obtain p-values by sampling normal random vectors with dependencies derived from observed patterns of variation in gene scores. This strategy yields correct type I errors and increases power. We identify a number of genes associated with hypertension in the Hutterites, including several previously identified in other populations.

Increasing Power To Identify Causal Variants Using Ascertainment For Targeted Resequencing

✦Michael Swartz, University of Texas School of Public Health, 1200 Pressler Street, Suite W920, Houston, TX 77030, *michael.d.swartz@uth.tmc.edu*; Cielito Reyes-Gibby, University of Texas M. D. Anderson Cancer Center; Bo Peng, University of Texas M D anderson Cancer Center; Sanjay Shete, University of Texas M. D. Anderson Cancer Center

Key Words: Ascertainment, Genome-Wide-Association Study, Causal Polymorphism, Targeted Resequencing

Researchers continue to use genome-wide association studies (GWAS) to find the genetic markers associated with disease. Recent studies have added to the typical two-stage analysis a third stage that uses targeted resequencing on a randomly selected subset of the cases to detect the causal single-nucleotide polymorphism (SNP). We propose a design for targeted resequencing that increases the power to detect the causal variant. The design features an ascertainment scheme wherein only those cases with the presence of a risk allele are selected for targeted resequencing. We simulated a disease with a single causal SNP to evaluate our method versus a targeted resequencing design using randomly selected individuals. The simulation studies showed that ascertaining individuals for the targeted resequencing can substantially increase the power to detect a causal SNP, without increasing the false-positive rate.

518 Issues in Adaptive Clinical Trials (I)

■

Biopharmaceutical Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Adaptive Blinded Sample Size Adjustment For Comparing Two Normal Means - A Mostly-Bayesian Approach

✦Andrew Montgomery Hartley, PPD, 929 N Front St, Wilmington, NC 28401, *khahstats@yahoo.com*

Key Words: clinical trials, adaptive designs, sample size re-estimation, bayesian analysis

Adaptive sample size re-determination (SSR) for clinical trials consists of examining early subsets of on-trial data, so as to adjust prior estimates of statistical parameters and sample size requirements. Blinded SSR, in particular, while in use already, seems poised to proliferate even further, due to recent draft guidance from the U.S. Food and Drug Administration. On the other hand, current blinded SSR methods offer little to no new information about the treatment effect (TE); the obvious resulting problem is that the TE estimate scientists might simply "plug in" to the SS formulae could be severely wrong. This paper proposes a blinded SSR method which formally synthesizes sample data with prior knowledge about the TE and the variance. It evaluates the method in terms of the average absolute deviation from the targeted statistical power, the type 1 error rate, the bias of the estimated TE and other measures. Under the conditions considered, the method reduces that average absolute deviation by roughly 15% to 25%, relative to another, established method. Simulations show the method to induce minimal bias and no increase to the type 1 error rate.

Proposal For A Two-Stage Adaptive Dose-Response Trial Design Using Repeated Measures Data

✦Anjela Tzontcheva, Merck Research Laboratories, , *anjela.tzontcheva@merck.com*; Susan Huyck, Merck Research Laboratories

Key Words: Dose-response characterization, repeated measures, simulation, two-stage design, adaptive design, constrained longitudinal data analysis

The chances of achieving successful late-stage drug development are greatly improved when the dose response has been definitively characterized allowing for the selection of the optimal dose for inclusion in confirmatory studies. Adequate characterization can be defined as identifying 4 dose levels where the lowest dose is not different than placebo, the 2 top doses plateau at the peak drug effect and the remaining dose is better than placebo, but lower than the plateau. Phase II dose-finding studies often represent the first large-scale evaluation of a new drug product in patients and adaptively designed studies provide opportunities for design adjustments that can occur because the true dose response range is unknown. This talk will present an adaptive dose response approach that improves the likelihood of adequate dose response characterization at the end of the study. Models are based on repeated measures data analyzed using a constrained longitudinal approach. Hypotheses first evaluated the presence of a dose response and then examine pairwise differences of drug vs. placebo. Simulations will be shown to demonstrate the operating characteristics.

A Three Part Adaptive Design- With Poc And Dose Finding For Induction Regimen And Dose Finding For Maintenance Regimen For An Immunology Trial

✦Usha Barai, Merck Research Laboratories, 2015 Galloping Hill Road, Kenilworth, NJ 07033-1300, *usha.barai@merck.com*

Key Words: Adaptive, POC, Induction Dose Finding, Maintenance Dose Finding, Futility, Dropping/Adding Doses

In certain disease area like Crohn's disease (CD), it is important to establish optimal induction as well as maintenance dose regimen. This three Part trial incorporates Proof of Concept (POC), Induction Dose Finding and Maintenance Dose Finding using adaptive features like futility, dropping/adding induction doses and optimizing maintenance doses based on modeling. We will discuss important features of the trial design including interim analyses with quantitative decision criteria based on the results of the simulation studies and practical issues in adaptive designs.

Adaptive Design For Bioequivalence Clinical Trials

✦Fengjuan Xuan, Merck Co. &, 2015 Galloping Hill Road, Kenilworth, NJ 07033 USA, *fengjuan.xuan2@merck.com*; Ferdous Gheyas, Merck Co. &

Key Words: Adaptive design, bioequivalence, conditional power, futility assessment, clinical trial

Adaptive design have been applied to superiority clinical studies to reduce drug development cost, while the application in bioequivalence studies is much less researched. Traditionally, a pilot study with a small sample size is conducted before a pivotal bioequivalence study to collect information such as variability and the GMR (geometric mean ratio) for comparing the experimental formulation with the reference formulation. This traditional drug development process is very time-consuming and costly. Therefore, adaptive bioequivalence study with two stages can significantly reduce the drug development cycle, and the cost of new drug. This presentation will discuss the design of a pivotal bioequivalence study where a futility assessment is performed at the interim analysis by an independent statistician . A closed-formula were derived to calculate the conditional power for repeated crossover design as well as other complex designs.

Adaptive Clinical Trial Design And Simulation Of A Phase 2 Migraine Headache Proof-Of-Concept Study Using Adaptive Design Explorer

✦Inna Perevozskaya, Pfizer, Inc., , *Inna.Perevozskaya@pfizer.com*; Jihao Zhou, Pfizer, Inc.; Pamela D. Garzone, Pfizer, Inc.; Phil Stanley, Pfizer, Ltd; John P. Huggins, Pfizer, Ltd; Gilbert Y. Wong, Pfizer, Inc.; Ronald G. Menton, Pfizer, Inc.

Key Words: adaptive design, experimental design, clinical trial, dose-response, proof of concept, migraine headache

As advocated by the Critical Path Initiative, adaptive designs have received a great deal of attention. The recent FDA guidance on adaptive clinical trials seems to favor their use for early phase trials where scientific rigor is blended with certain regulatory flexibility. However, lack of adaptive trial design software poses a major challenge for their implementation in the industry. Here we report some of our recent experiences in building up adaptive designs and simulations for an early clinical study for a migraine headache compound using the software Adaptive Design Explorer. The primary efficacy endpoint for this study is the change from baseline in the frequency of migraine periods over 28 days. Several adaptive design types utilizing various dose-response models (e.g. logistic and Emax model functions, non-linear dynamic model (NDLM), etc.) are chosen and compared to select the best performing design. Number of interim analyses and other design parameters are optimized as well. The best performing designs are then compared to a traditional parallel group design in order to quantify the efficiency of adaptive designs.

Interval Estimation In Two-Stage Clinical Trials That Are Stopped At The First Stage

Samuel S. Wu, University of Florida Department of Biostatistics; ✦Dan Neal, University of Florida Department of Biostatistics, 1329 SW 16th Street Room 5231, Gainesville, FL 32601, *dwn@biostat.ufl.edu*

Key Words: adaptive design, drop-the-losers, two-stage, confidence limits, clinical trial design

The increasing use of adaptive clinical trial designs has generated interest in methods to establish a lower confidence limit for the difference in effect between treatment and control. Recent papers have explored interval estimation of effect size in two-stage drop-the-losers trials; however, the existing methods are conservative when the trial is stopped at the first stage. This is particularly undesirable when the trial is stopped for clear superiority because the effect size might be large. Similar to Chang (1989) and Tsiatis et al. (1984), this paper will develop procedures for constructing better lower confidence limits in this situation.

Martingale-Based Estimators Following Group Sequential/Adaptive Designs

Key Words: Group Sequential Design, Adaptive Design, Martingale, Clinical Trial

In the context of multi-stage (group sequential or adaptive) designs, the naive estimators,which equal the difference of the sample means between the treatment groups up to the stopping stage, are biased. For group sequential designs, two common approaches were used to adjust the bias explicitly or get unbiased estimators from sufficient statistics. However, these methods either apply to a restricted set of group sequential designs or are inapplicable to adaptive designs. We now propose a brand new series of estimators which are solutions of equations based on Martingale processes; meanwhile, another series of estimators which combine the advantages of naive estimators and martingale-based estimators are also proposed. Our approaches can be easily extended to the adaptive designs and confidence intervals can also be obtained. These approaches show high accuracy and flexibility in both simulations and a case study from pharmaceutical industry.

519 Topics in Group Sequential Designs and Interim Analysis ■

Biopharmaceutical Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Software Applications To Group Sequential Trial Design And Monitoring

✦Li Zhu, Amgen Inc., One Amgen Center Drive, Mail Stop 24-2-A, Thousand Oaks, CA 91320, *zhul@amgen.com*; Liyun Ni, Amgen Inc.; Bin Yao, Amgen

Key Words: Conditional power, Error spending function, Interim monitoring, Stopping boundary

Group sequential methods are commonly used in clinical trials as they allow valid inference when statistical tests are performed on accumulating data from ongoing trials. Several software applications have become available to meet the computational need for group sequential trial design and monitoring. This talk will review several up-to-date software applications. The capabilities and special features of the software applications are illustrated and compared through practical examples.

Considerations On Decision Rules For Futility Based On Multiple Studies

✦Annpey Pong, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, NJ 07065, *annpey.pong@merck.com*; Jun Zhao, Merck & Co., Inc.

Key Words: Interim analysis, possibility of success, futility, program-wise futility

In drug research and development, interim analyses may be applied to evaluate the possibility of success in order to stop a trial early for futility. According to literature, the futility decision (GO/NOGO) is mostly applied to an individual clinical trial. However, to reduce the lengthy drug development, it is common to have multiple ongoing studies in a development program at the same time for a new indication; and it is also common to have failed or negative studies in the program. The goal of the futility analysis is then extended to make GO or NOGO decision on whether to continue developing the entire program. In this presentation, the decision rules based on various considerations on multiple studies are proposed. In addition, example is given from a central nervous system drug development program.

Optimal Designs With Futility Stopping For Diagnostic Imaging Studies

✦Aida Aydemir, Bayer HealthCare Pharmaceuticals Inc., 340 Changebridge Road, Montville, NJ 07045, *aida.aydemir@bayer.com*; Igor Zurbenko, State University of New York at Albany; Daniel Haverstock, Bayer HealthCare Pharmaceuticals Inc.; Christoph Gerlinger , Bayer Schering Pharma AG

Key Words: Futility, Conditional power, Futility treshold, Interim analysis, Current trend, Confidence intervals

For early phase drug development studies early stopping for futility is essential to minimize efforts on non performing products. For diagnostic imaging studies where specificity and sensitivity of the testing procedure have to be analyzed simultaneously this is even more crucial. Due to the restrictions of our planned trials we restricted our work to designs with only one interim analysis. The aim was to find the optimal futility threshold, interim time point (and sample size) for the interim analysis as well as the optimal allocation of fl, the error of the second kind. We adapted several sequential designs suggested in the literature to our problem of Futility Stopping for Diagnostic Imaging Studies. We investigated their operating characteristics in a large simulation study. A practical application of a recent diagnostic imaging study is given.

Interim Look At Subgroups In Clinical Trials

✦Chengxing Lu, Merck Research Labs, 351 N Sumneytown Pike, North Wales, PA 19454 US, *chengxing_lu@merck.com*; Anthony Rodgers, Merck Research Labs; Mani Lakshminarayanan, Merck & Co

Key Words: Conditional power, Bayesian predictive power, Subgroup analysis, Interim look

When various subgroups of patients respond to the treatment dose regimen in different patterns, the trial teams tend to investigate the efficacy in subgroups at interim stage in addition to the primary population. The probability of observing a statistically significant treatment effect at the end of a trial based on the overall interim data and the subgroups can be studied using conditional power and Bayesian predictive power. One issue addressed in this talk is how to separate out the criteria of observing the conditional power between subgroups of patients in decision making. Further discussion will focus on the determination of the response rate that needs to be seen at the interim which associates to a probability threshold (conditional power or predictive power) to facilitate decision making. Both frequentist and Bayesian methods in interim monitoring of subgroups will be studied. Simulation results will also be presented in support of the comparison.

Group Sequential Tests Of Futility: The Relationship Between Type Ii Error Spending And Conditional Power Approaches

✦Thomas W Dobbins, Merck & Co., Inc., 351 N. Sumneytown Pike, UG1D-80, North Wales, PA 19454-2505, *thomas_dobbins@merck.com*

Key Words: Adaptive Design, Conditional Power, Futility Boundaries, Group Sequential Test, Type II Error Spending, Stochastic Curtailment

In a clinical trial incorporating a group sequential test that allows early stopping for efficacy or futility (GSTEF), or early stopping for futility only (GSTF), two common approaches to testing for futility are Type II error spending and conditional power. As their names imply, these approaches differ in how the probability of incorrectly stopping for futility is specified; consequently, the statistical properties of the two approaches are not easily compared. We utilize a partitioning of the Type II error derived by Dobbins (2011) to obtain the exact relationship between conditional power and Type II error spend for any K-group GSTEF. In particular, the conditional power approach can be completely specified as a GSTEF using Type II spending and non-binding futility boundaries. We describe a simple procedure to obtain the Type II error spending sequence for the interim tests of futility directly from the futility boundary corresponding to a specified conditional power, and vice versa, and we consider an example that illustrates the correspondence.

Futility Analysis Approaches In Clinical Equivalence Trials

✦Christine K. Gause, Merck Research Laboroatories, UG-1CD38, 351 N. Sumneytown Pike, North Wales, PA 19454, *christine_gause@merck.com*

Key Words: equivalence, futility, group sequential

In many studies with an active comparator, a formal demonstration of non-inferiority is the goal, and a variety of methods exist for demonstration of non-inferiority using synthesis approaches. These methods rely on an assumption that the active comparator effect is the same in the current trial as in previous historic trials. We investigate synthesis approaches applied to the clinical equivalence setting, with incorporation of an interim futility analysis to potentially stop a trial if it is unlikely that the experimental drug is equivalent to the active comparator. Both group sequential spending functions and Bayesian approaches to equivalence testing that incorporate prior information for the active comparator will be considered. A simulation study is conducted to compare the approaches under a variety of conditions for a trial with a binary outcome.

Blind Review Of Planned Analyses For Clinical Trials

✦Lan Liu, Quintiles Inc, 6700 W 138th Ter Apt 2012, Overland Park, KS 66223, *Lan.Liu@quintiles.com*; John Schollenberger, Quintiles Inc

Key Words: Blind review, Planned analysis

The ICH E9 guidance suggests that a review of the data be done after trial completion and prior to treatment unblinding, for the purpose of finalizing the planned analysis-in particular, the choice of a statistical model. However, for Phase III clinical trials especially, the model chosen typically has been used in previous analyses, and assumption checking has already been performed. Thus, at this stage, it is usually not necessary, for example, to check the need for a transformation. Further, sensitivity analyses to show the robustness of the planned analysis to the method of handling missing values are typically already included in the analysis plan. In this situation, what types of checks should be included in the blind review? The guidance provides some suggestions-e.g., a check of the possible need to include additional covariates in the model, and a reconsideration of the use of parametric vs. nonparametric methods. But, what other types of checks should be done, and how? Should it be driven by significance tests? A suggested framework for the blind data review, focusing on the statistical modeling, will be discussed.

520 Modeling of Complex Correlated Data and Images ■●

ENAR, International Indian Statistical Association, Section on Physical and Engineering Sciences

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Empirical Null Distribution Modeling For Signal Detection In Fmri

✦Shuzhen Li, University of Minnesota, 1039 29th AVE SE APT F, Minneapolis, MN 55414, *lixxx466@umn.edu*; Lynn Eberly, University of Minnesota

Key Words: fMRI, FDR, Empirical null distribution

In signal detection of fMRI data, the inherent neural and spatial correlation among voxels could change the theoretical distribution of test statistics across voxels and thus compromise the accuracy and efficiency of region of interest (ROI) detection. In this article, a statistical parametric map (SPM)with moderated t statistics (Smyth, 2004) is created by considering a stabilized variance estimate instead of the usual sample variance estimate for typical fMRI data with small sample size compared to the large number of voxels. The empirical null distribution of the moderated t statistics, an assumed t distribution with scale parameter s and non-central parameter ?, is estimated through a maximum likelihood approach based on truncated data. A Poisson regression model is also implemented to estimate the nonparametric marginal distribution of all voxels. False discovery rate (FDR) thresholding can then be applied. We consider methods of Benjamini and Hochberg (1995) and computed local FDR (fdr, Efron 2007b) based on both regular and moderated t statistics. The criteria of true FDR and sensitivity are compared to demonstrate the performance of our methods in large simulations.

Generalized Estimating Equation And Bootstrapping Approaches To Evaluate Functional Connectivity

✦Gina M D'Angelo, Washington University, 660 South Euclid Avenue, Box 8067, St. Louis, MO 63110-1093 USA, *gina@wubios.wustl.edu*; Nicole Lazar, Department of Statistics; Guofu Zhou, Washington University

Key Words: fMRI, GEE, transition model, bootstrap, Functional connectivity

In the neurodegenerative disease area it is often of interest to evaluate relationships between regional brain measures. Specifically, in resting-state functional magnetic resonance imaging (fMRI) studies an objective is to assess inter-regional correlations at resting-state with no task or stimulus. Resting-state fMRI studies produce cross-sectional time-series data across many regions of the brain. We propose using a generalized estimating equation (GEE) transition model to account for the lag and model the within-subject correlation for each region. In addition, the residuals from this GEE transition model will be extracted to correlate brain regions and assess between group differences. The GEE approach and standard pooling approach of group averages of the fisher-z transformation will be compared with simulation studies to determine statistical properties. Various bootstrapping approaches, including the block bootstrap and circular bootstrap, will be applied to estimate the standard error and confidence intervals. Both methods will be demonstrated with an Alzheimer's disease connectivity study in a population of Alzheimer's disease subjects and healthy control subjects.

Change Point Detection In Fmri Functional Connectivity Networks

✦Lucy Robinson, Johns Hopkins University, 3400 n charles st, 100 Whitehead Hall, Baltimore, MD 21218, *lucy.f.robinson@gmail.com*; Lauren Atlas, Columbia University; Carey Priebe, Johns Hopkins University

Key Words: fMRI, networks, change point, neuroscience

We propose a new technique for detecting time-dependent structure in networks of brain regions. Functional connectivity networks are derived from fMRI data, using partial coherence to describe the degree of connection between spatially disjoint locations in the brain. Existing techniques for functional connectivity networks typically assume the network is static over time. Our goal is to determine whether brain network topology remains stationary, or if subsets of the network exhibit changes over unknown time intervals. Changes in network structure may be related to shifts in neurological state, such as those associated with learning, drug uptake or experimental stimulus. Our approach can detect localized changes, allowing us to infer which nodes in a network exhibit altered connectivity behavior. Using spectral clustering, we define a time-dependent block structure and a test for temporal homogeneity based on a modified likelihood ratio. Data from an experiment studying mediation of pain response by expectancy of relief are analyzed.

Measurement Error In General Multivariate Linear Model

✦Tielin Qin, Emory University, 1518 Clifton Road NE, ATLANTA, GA 30322 US, *tqin@emory.edu*; Vicki S. Hertzberg, Emory University

Key Words: Measurement error, pseudo MLE, EM algorithm

We are motivated by the need to modeling the true count of white blood cells in the Protective Immunity Project (PIP) study conducted at the Emory Transplant Center. The recorded data obtained from flow cytometer may contain measurement errors. General multivariate linear model are used, and the measurement errors are presented in the dependent variables only. The parameters in the measurement error model are estimated by using independent, external calibration data. The pseudo-maximum likelihood estimators and their asymptotic properties are developed. A simulation study is conducted to compare the performance of the pseudo MLE approach and the simply adjusted/imputed data approach.

Bayesian Modeling Of The Dependence In Longitudinal Data Via Partial Autocorrelations And Marginal Variances

✦YANPIN WANG, UNIVERSITY OF FLORIDA, UNIVERSITY OF FLORIDA, GAINESVILLE, FL 32611, *yanpin@stat.ufl.edu*; Michael Daniels, University of Florida

Key Words: Bayesian Modeling, Lingitudinal Data, Partial Autocorrelation, Markov Chain Monte Carlo, Generalized Linear Model

Two major obstacles are high dimensionality and positive-definiteness as we estimate and model a correlation matrix for longitudinal data. In addition, incorrectly modeling the correlation matrix often results in bias in estimating mean regression parameters as missing data exists. In this work, we introduce regression models for partial autocorrelations using Fisher's z-transform as the link function. The partial autocorrelations proposed can freely vary in the interval (-1, 1) while preserving positive definiteness of the correlation matrix is a key factor of our work. We propose a class of priors for the regression coefficients of the transformed partial autocorrelations and examine their behavior via simulations. The approach is illustrated on data from a pharmacological treatment of schizophrenia clinical trial.

Empirical-Bayesian Inference For Count Data Using The Spatial Random Effects Model

✦Aritra Sengupta, Ohio State University, 1958 Neil Avenue, Cockins Hall, Room 404, Columbus, OH 43210 United States, *sengupta.11@osu.edu*; Noel Cressie, Ohio State University

Key Words: Poisson model, geostatistical process, maximum-likelihood, EM algorithm, method-of-moments, MCMC

This paper is concerned with inference for spatial data in the form of counts. We consider a Poisson model for the counts, and assume an underlying geostatistical process for the mean of the Poisson distribution. We develop maximum-likelihood estimates for the parameters of the continuous process using the expectation-maximization (EM) type algorithm. The starting value for the EM algorithm is critical, and we obtain our starting values using method-of-moment estimators. The expectations in the E-step of the EM algorithm are not available in closed form, so we use some numerical and theoretical approximations to the expectations required in the E-step. Empirical-Bayesian inference, based on an MCMC, will then be discussed.

Estimating Equations For Regression Models With Cluster-Specific Intercepts

✦Kunthel By, Dept of Biostatistics, UNC, Chapel Hill, NC 27599, *kby@email.unc.edu*; Bahjat Qaqish, Dept of Biostatistics

Key Words: biased sampling, correlated data, GEE, nuisance parameters

GEE is not feasible for marginal models having cluster-specific intercepts. Imposing a mixing distribution on the intercepts offers a possible solution provided that generalized linear mixed model assumptions are satisfied. When these assumptions are not met, parameter estimates are generally biased. A simple procedure for constructing estimating equations is proposed that enable consistent estimation of parameters associated with cluster-varying covariates and is applicable regardless of whether cluster-specific intercepts are treated as fixed or random. Connections to conditional likelihoods and the Cox model are discussed. An application to outcome-dependent sampling based on cluster totals is proposed; knowledge of the sampling rates is not required. We show how existing software can be used to implement our estimating equations.

521 Models for Binary and Count Data

Business and Economic Statistics Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Doubly-Inflated Poisson (Dip) And Related Regression Models

✦Manasi Sheth, Old Dominion University, 1405 Dure Road, Unit 1C, Norfolk, VA 23502, *msheth@odu.edu*; Dr. N. Rao Chaganty, Old Dominion University

Key Words: Count Data, Doubly Inflated Poisson, Maximum Likelihood, Method of Moments, Relative Efficiency

Usually over-dispersed count data consists of certain values occurring more frequently than allowed by the common parametric families of distributions. For dealing with zero-inflated count data with covariates, Lambert (Technometrics, 1992, pp. 1 - 14) proposed the zero-inflated Poisson (ZIP) regression model. In this talk, we introduce Doubly Inflated Poisson (DIP) models for count data situations where there is another inflated value $k > 0$. We present the distributional properties of these models. For the data consisting of un-grouped as well as grouped frequencies, with and without covariates, we discuss parameter estimation using maximum likelihood (ML) and method of moments. Asymptotic and small sample comparisons show that the ML estimators are far superior than the moment estimators. We illustrate the application of DIP models to analyze length of stay (LOS) in the hospital for patients with pancreatic disorder. Parameter estimation and analysis of count data with a negative binomial data-generating process will also be presented using a Doubly Inflated Negative Binomial (DINB) Model.

Flexible Bivariate Count Data Regression Models

✦Shiferaw Gurmu, Georgia State University, Department of Economics, P.O.Box 3990, Atlanta, GA 30033, *sgurmu@gsu.edu*; John Elder, Colorado State University

Key Words: Series estimation, Negative correlation, Unobserved heterogeneity, Multivariate counts, Tobacco use, Health care utilization

The paper develops semiparametric estimation methods for bivariate count data regression models. We develop series expansion approach in which dependence between count variables is introduced by means of stochastically related unobserved heterogeneity components, and in which, unlike existing commonly used models, positive as well as negative correlations are allowed. In implementation, we use bivariate expansions based on the generalized Laguerre polynomials. Extensions that accommodate for excess zeros, truncated and censored data and multivariate generalizations are also given. The first application examines the socio-economic and demographic determinants of tobacco use in the context of the joint modeling of the daily number of smoking tobacco and number of chewing tobacco based on household survey data. We also analyze jointly two health utilization measures, number of consultations with a doctor and non-doctor consultations. One of the key contributions is in obtaining a computationally tractable closed form of the model with flexible correlation structure. Results from the applications and simulation experiments confirm that the method is feasible and perform well.

A Semiparametric Approach For Multivariate Longitudinal Count Data

✦Darcy Steeg Morris, Cornell University, 171 East State Street, Apartment 311, Ithaca, NY 14850, *dsteeg@alumni.princeton.edu*

Key Words: Generalized Estimating Equations, Correlated Count Data, Generalized Linear Mixed Models, Unobserved Heterogeneity, Longitudinal Data Analysis

A semiparametric method for estimating the marginal regression and association parameters in a multivariate longitudinal random effects count model is considered. Using the generalized estimating equation (GEE) framework, a specific form of the covariance matrix of the response vector is imposed which induces dependence over time and outcomes using random effects. This moment based method reduces the computational burden associated with a high-dimensional joint distribution by avoiding parametric assumptions on the response and unobserved effects. Through a simulation study, performance of the semiparametric estimators is compared with estimators from a pairwise likelihood approach. Both of these methods are then used to analyze a dataset of insurance claim counts for three types of coverage over time.

On Binomial AR(1) Models

✦Yunwei Cui, University of Houston-Downtown, Department of Computer and Mathematical Sciences, One Main Street S705, Houston, TX 77002, *cuiy@uhd.edu*

Key Words: Integer-valued time series

Integer-valued time series arise in many practical settings. A renewal process formulation of the problem is introduced. The model satisfies an AR(1) recursion in cases where the renewal lifetime has a constant hazard rate beyond lag 1. The explicit asymptotic variances of the conditional least squares estimators and the maximum likelihood estimators are derived.

On The Autopersistence Functions And The Autopersistence Graphs Of Binary Autoregressive Time Series

✦Chao Wang, The University of Hong Kong, Room 518, Meng Wah Complex, The University of Hong Kong, Hong Kong, China, *chaowang@hku.hk*; Wai Keung Li, University of Hong Kong

Key Words: autopersistence function, autopersistence graph, binary time series, Markov chain, consistency, asymptotic normality

The classical autocorrelation function may not be an effective and informative means in revealing the dependence features of a binary time series, which, without loss of generality, is assumed to take values 0 and 1. Recently, the autopersistence functions, defined as the functions of the lag k with the value being the conditional probability with the observation at time t being 1 given the particular observation at time t-k, have been proposed as alternatives to the autocorrelation function for binary time series. In this article we consider the theoretical autopersistence functions and their natural sample analogues, the autopersistence graphs, under a binary autoregressive model framework. Some properties of the autopersistence functions and the asymptotic properties of the autopersistence graphs are discussed. The results have potential application in the modelling of binary time series.

On The Economic Meaning Of Interaction Term Coefficients In Non-Linear Binary Response Regression Models

✦Andrew F. Siegel, University of Washington, Foster School of Business, Box 353226, Seattle, WA 98195-3226, *asiegel@uw.edu*; Adam C. Kolasinski, University of Washington

Key Words: Logit regression, Probit regression, Categorical data analysis, Interactive effects, Estimation techniques

We show that it is perfectly correct to use just the interaction term, along with its standard error, to draw inferences about interactive effects in binary response regression models. This point is currently in dispute among applied econometricians, some of whom insist that simply relying on the interaction term is incorrect, since the cross partial derivative of the probability of occurrence with respect to interacted covariates can, for some observations, have the sign opposite to that of the interaction term coefficient. We show that this sign flip results from a mechanical saturation effect that is of no importance to researchers who recognize that small changes in probability are more important near the boundaries than near the center. For such researchers, the interaction term coefficient provides a more meaningful measure of interactive effects than does the cross partial derivative of the probability itself. We introduce an alternative cross partial derivative of the probability for which these sign changes cannot occur. Finally, we demonstrate some simple and intuitive ways of interpreting the economic meaning of interaction term coefficients.

Should We Always Use The Logarithmic Transformation For Positive Response Variables?

✦Shengwu Shang, Department of Economics,Michigan State University, 110 Marshall-Adams Hall, Michigan State University, East Lansing, MI 48824-1038, *shangshe@msu.edu*; JEFFREY M. WOOLDRIDGE, Department of Economics,Michigan State University

Key Words: LFE;, PQML;, GMM;, Consistency, Efficiency;

ABSTRACT We compare three main estimation methods for positive response variable-- FE method for log linear model (LFE), Poisson Quasi-Maximum Likelihood (PQML) and Generalized Method of Moment (GMM) -- by Mont Carlo Simulation and real life data set. It is not surprising that LFE estimator is not consistent when PQML is; however, we do find circumstance where both LFE and PQML estimators are consistent plus LFE is more efficient. With this regard, we introduce GMM to improve the efficiency of PQML estimator as well as keeping the consistency; this way also finds a solution to the problem raised in Wooldridge (1999). From the simulation results, we find that GMM can reduce the standard errors of PQML estimators by almost a half. We also apply the GMM to a real life data set and the result shows that GMM is applicable.

522 Time Series and Network ●

IMS

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Scalable Mixture Sequential Multi-Sensor Change-Point Detection Procedure

✦Yao Xie, Stanford University, 109 McFarland CT, APT 400, Stanford, CA 94305 USA, *yaoxie@stanford.edu*; David Siegmund, Stanford University

Key Words: Multisensor detection, Change-point problem, Sequential detection, False-alarm rate, generalized likelihood ratio (GLR), maximum of ranom field

We develop a scalable mixture (SaM) procedure for unstructured multi-sensor sequential change-point detection problems, where multiple sensors are distributed to monitor the emergence of a signal that changes the means of observations from part of sensors. We assume that the affected sensors and their post-change means are unknown. SaM uses an estimate for the fraction of affected sensors, p, and sums the nonlinearly transformed generalized likelihood ratio statistics (GLR) from each sensor. The sum is then compared with a threshold chosen to maintain a certain false alarm rate. In SaM, the nonlinear transform achieves noise suppression from unaffected sensors by a soft-thresholding, which is controlled by p. We derive a closed-form false alarm rate for SaM as well as its expected detection delay with good accuracy verified numerically. We also show that SaM does not require an accurate estimate of p. Finally, we demonstrate using numerical studies that for a given false alarm rate, SaM has smaller expected detection delay than other procedures for unstructured problems such as taking the maximum or sum over the local GLR statistics, and the recently proposed Mei's procedure.

Modeling And Forecasting The Time Series Of Treasury Bond Yield Curves

✦Cong Feng, Dept, of Statistics, University of Georgia, Statistical Building, University of Georgia, Athens, GA 30602 U.S, *congfeng@uga.edu*

Key Words: functional time series, eigenfunctions principal component analysis, principal component analysis, spline smoothing

A novel method is proposed for forecasting the time series of smooth curves, using functional principal component analysis in combination with time series modeling and forecasting the scores. In this way, we can achieve the smoothing, dimension reduction and prediction at the same time with the expedient computation. The research problem is motivated by the demand to forecast the time series of economic functions, such as Treasury bond yield curves. Extensive simulation studies have been carried out to compare the prediction accuracy of our method with other competitor's methods. The proposed methodology is applied to forecasting the yield curves of UK government bond. It is the jonit work with Li Wang and Lynne Seymour.

The Fundamental Theorem Of Asset Pricing With No Short Selling

✦Scott Mcclintock, West Chester University, 25 University Ave, RM 121, West Chester, PA 19383, *smcclintoc@wcupa.edu*; Stephen Clark, Emory University

Key Words: arbitrage, martingale, applied probability

The Fundamental Theorem of Asset Pricing, under the assumption of restricted short selling, states that there is no feasible arbitrage if and only if there exists an equivalent supermartingale measure. Such a result has been demonstrated by authors such as Schurger (1996) and Evstigneev, Schurger, and Taksar (2004). Borrowing on ideas used by Kabanov and Kramkov (1994) and Schurger (1996), and under the assumption of a degeneracy condition, we provide a simple, elegant proof of this result that uses only basic probability arguments and techniques. We then go on to sharpen the preceding result and demonstrate that, assuming this degeneracy condition holds, that a market satis?es no feasible arbitrage if and only there exists an equivalent martingale measure.

A Bayesian Adaptive Singular Control Problem With Discretionary Stopping Arising From Finance

✦Stephane Villeneuve, UniversitÈ de Toulouse 1 capitole, 21 allÈe de brienne, Toulouse, 31000 France, *stephane.villeneuve@univ-tlse1.fr*

Key Words: Singular control, optimal stopping, Skohorod reflection, filtering, finance

Stochastic optimization problems that combine features of both bounded variation control and stopping are relatively scarce in the applied probability literature. Recently, mixed stochastic control problems have emerged from corporate finance in continuous time. Bounded variation control problems arise in the corporate finance literature when we consider the optimal liquidity management (dividend and issuance policies) of a firm subject to costly external financing. On the other hand, discretionary stopping in stochastic control arises naturally in solvency risks problems where the shareholders have to determine endogenously the optimal time to liquidate a project. In this paper, we analyse the interaction between dividend policy (the bounded variation control) and optimal abandonment (discretionary stopping) of a liquidity constrained firm whose profitability is partially known. This lead us to study a mixed singular/optimal stopping problem under partial information on the drift of a diffusion that we solve explicitly in some special cases. Moreover, sufficient conditions for the existence of optimal strategies relying on Skohorod reflection are provided.

Statistical Image Analysis And Percolation Theory

✦Mikhail Langovoy, Max Planck Institute for Biological Cybernetics, Max Planck Institute for Developmental Biology, Tuebingen, Germany, , *mikhail.langovoy@tuebingen.mpg.de*; Michael Habeck, Max Planck Institute for Biological Cybernetics and Max Planck Institute for Developmental Biology; Bernhard Schoelkopf, Max Planck Institute for Biological Cybernetics

e develop a novel method for detection of signals and reconstruction of images in the presence of random noise. The method uses results from percolation theory. We specifically address the problem of detection of multiple objects of unknown shapes in the case of nonparametric noise. The noise density is unknown and can be heavy-tailed. The objects of interest have unknown varying intensities. No boundary shape constraints are imposed on the objects, only a set of weak bulk conditions is required. We view the object detection problem as hypothesis testing for discrete statistical inverse problems. We present an algorithm that allows to detect greyscale objects of various shapes in noisy images. We prove results on consistency and algorithmic complexity of our procedures. Applications to cryo-electron microscopy are presented.

Multiscale Diffusion Approximations For Stochastic Networks In Heavy Traffic

✦Xin Liu, University of North Carolina at Chapel Hill, NC 27599, *xinliu@unc.edu*; Amarjit Budhiraja, University of North Carolina at Chapel Hill

Key Words: Diffusion approximations, Queueing networks in random environment, Heavy traffic, Multiscale analysis, Reflected Markov modulated diffusions, Constrained martingale problems

Open queueing networks with time varying arrival and service rates and routing structure are studied. Time variations are governed, in addition to the state of the system, by two independent Markov processes X and Y. Transition times of X are significantly smaller than typical interarrival and service times whereas the reverse is true for the Markov process Y. By a suitable martingale problem, diffusion approximations for such multiscale systems are established under a suitable heavy traffic condition. It is shown that properly normalized queue length processes converge weakly to a reflected diffusion with drift and diffusion coefficients that, in addition to depending on the state process, are functions of the invariant distributions of X. Additionally, the coefficients are modulated by a finite state Markov process which is independent of the driving Brownian motion

Proposed Conditional Probability Statistical Model For Recurrent Events With Application To Hypoglycemia Data In Diabetes Clinical Trials

✦Xiaodan Wei, sanofi-aventis, 200 crossing blvd, Bridgewater, NJ 07921, *xiaodan.wei@sanofi-aventis.com*

Key Words: Recurrent event, diabetes, conditional probability

We propose a statistical model handling binary data at each time point and allowing the certain correlation structure between time points based on the conditional probability of previous event. Simulation results will be provided to compare various models on Empirical Significance Level (ESL) and power performance of fitting recurrent events data using different testing method, including: Poisson, Negative binomial model with dispersion parameter, zero-inflated Poisson, zero-inflated Negative binomial, Poisson GEE, CMF (cumulative mean function) using robust variance estimate, CMF with previous event information as covariate, using robust variance estimate and our proposed method. Then we discuss the advantages and disadvantages of using each method on event count or event rate with application to the hypoglycemia events from diabetes clinical trials. Finally we provide recommendations on the most suitable statistical model and testing method to use for hypoglycemia data.

523 Econometric Methods and Applications ■●

Business and Economic Statistics Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Inflation Expectations and Risk Premium

✦Luis Fernando Melo, Central Bank of Colombia, Carrera 7 # 14-78, Bogota, ----- Colombia, *lmelovel@banrep.gov.co*

Key Words: Break Even Inflation, Inflation risk premium, Inflation expectations, State space models, Affine term structure models

The Break Even Inflation is estimated using the nominal and real government Colombian bonds for the period January 2003 to November 2009. This measure is decomposed in inflation expectations and inflation risk premium. The inflation expectations are calculated using a state space representation of an extended affine term structure model. In order to improve the forecasts, this model incorporates the inflation expectations 12 months ahead of the Colombian Central Bank survey. The results show an inflation expectation downward trend, which may be related to an increasing confidence in monetary policy. This hypothesis is also supported by a decreasing trend in the inflation risk premium for medium and long term maturities (two and five years). Finally, the results indicate that the break even inflation is a good indicator of the inflation expectations for short term forecast horizons (one year).

Frame Coverage Error Estimation In The 2009 Residential Energy Consumption Survey

✦Edgardo Cureg, U.S. Energy Information Administration, 1000 Independence Ave SW, Washington, DC 20585, *edgardo.cureg@eia.gov*

Key Words: frame coverage error, address-based sampling, non-response

For the first time in its 30-year history, the 2009 round of the Residential Energy Consumption Survey (RECS) relied on residential addresses from the U.S. Postal Service's Delivery Sequence File (DSF) to construct a significant portion of its housing unit frame. In this study we examine the coverage error of this frame and its relationship to the underlying non-response mechanism, their combined effects on the survey results, and potential remedies to remove or adjust for these effects.

A Semiparametric Interest Rate Model Based On Reducible Stochastic Differential Equations And Pseudo Maximum Likelihood Estimation

✦Ruijun Bu, The University of Liverpool, , *ruijunbu@liv.ac.uk*; Kaddour Hadri, Queen's University Belfast

Key Words: Semi-parametric Models, Non-parametric Estimation, Multivariate Interest Rate Models, Reducible Stochastic Differential Equations, Maximum Likelihood Estimation, Time-Varying Copulas

We propose a new semiparametric model for interest rate processes based on Reducible Stochastic Differential Equations (RSDEs). The idea of using RSDEs for modelling interest rates was pioneered by Bu et al (2011) who proposed two classes of nonlinear SDEs that are reducible to either Ornstein-Uhlenbeck or Cox, Ingersoll, and Ross (1985) process via parametric transformations. In this paper, we extend Bu et al (2011) by allowing the transformation function to be completely unspecified. In contrast to existing semiparametric models (e.g. AÔt-Sahalia 1996, Kristensen 2010) where either the drift or the diffusion must be completely parametric, both of them are semi-parametric in our framework. Their shapes are thus not limited to the imposed parametric structure. An explicit semiparametric functional estimate of the transition density function is derived. We therefore propose to estimate the model by pseudo maximum likelihood. We further extend our model to multivariate case by the copula approach. In our empirical application, we study the goodness of fit of the new model to UK and US short rates and examine their conditional dependence structure in our multivariate framework.

A Class Of Exact F-Tests For Capital Asset Pricing Models

✦Jiajuan Liang, University of New Haven, 300 Boston Post Road, College of Business, West Haven, CT 06516, *jliang@newhaven.edu*

Key Words: Capital asset pricing, Dimension reduction, Exact F-test, Monte Carlo method, Portfolio efficiency

The tradeoff between risk and expected return in the formation of an investment portfolio is a central focus of modern financial theory. The Capital Asset Pricing Model has been widely employed to study the efficiency of a portfolio. Testing whether a portfolio is efficient was proved to reduce to testing a statistical hypothesis on a parameter constraint in a multivariate linear regression model. In this article we develop a class of F-tests to evaluate the portfolio efficiency and compare the new tests with existing ones in the case of high dimension with small sample sizes by using the Monte Carlo method. Empirical results show that the new tests outperform the classical ones in the circumstances where the number of observation time points is close to the number of assets

(securities) in the portfolio and they still work effectively even if the number of observation time points is less than the number of assets (securities) in the portfolio. We illustrate the result by using real data and test the circumstances where the number of data points is small and classical tests have no power.

A Monte Study On A Class Of Tests For Capital Asset Pricing Models

✦Luoyuan Liang, Illinois Institute of Technology, 80 Deepwood Drive, Madison, CT 06443, *yoyo.luoyuan@gmail.com*

Key Words: Capital asset pricing model, Dimension reduction, Generalized F-test, Monte Carlo study

The Capital Asset Pricing Model (CAPM) has been widely employed to study the efficiency of an investment portfolio in modern financial theory. Many statistical tests for this purpose have been proposed in the literature. Existing Monte Carlo studies show that all proposed tests possess both benefits and limitations in testing CAPM. In this paper, we carry out a Monte Carlo study on the empirical performance of a class of dimension-reduction-type F-tests and a class of generalized F-tests for CAPM in the cases: 1) the sample size is much larger than the dimension; 2)the sample size is closer to the dimension; and 3) the sample size is smaller than the dimension. A simple comparison with some existing tests is illustrated by some selected CAPMs and a practical data set.

Empirical Evaluation Of Global Vector Autoregressions

✦Neil R Ericsson, Federal Reserve Board, Stop 20, 2000 C Street NW, Washington, DC 20551, *ericsson@frb.gov*

Key Words: cointegration, data aggregation, exogeneity, model evaluation, parameter constancy, VAR

Global vector autoregressions (GVARs) have several attractive features: a standardized economically appealing choice of variables for each country or region examined, a systematic treatment of long-run properties through cointegration analysis, and flexible dynamic specification through vector error correction modeling. Dees, di Mauro, Pesaran, and Smith (2007) develop a paradigm GVAR with 26 countries. The current paper empirically evaluates that GVAR with new tests from Ericsson (2010), focusing on exogeneity assumptions, parameter constancy, and data aggregation. The results indicate substantial room for an improved, more robust specification of that GVAR, with some tests suggestive of how to achieve such improvements.

524 Computing and the First Statistics Course ■●

Section on Statistical Education

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Is The T-Test Really Dead? Part 1

✦Richard De Veaux, Williams College, 28 Southworth ST, Williamstownm, MA 01267, *deveaux@williams.edu*; Paul Velleman, Cornell University

Key Words: Gossett, t-test, resampling, introductory course

Student (W. S. Gosset) showed that we only needed the mean, the standard deviation, and the sample size to describe the sampling distribution of the sample mean, with the additional assumption that the population was Normal and a sample size large enough to justify the asymptotic result. Gosset's result is often at the center of our introductory courses. But have we learned the wrong lesson? Gosset derived his sampling distribution by simulating. Using 3000 observations that seemed to be Normal, he drew 750 samples of size 4, computed their means and standard deviations with a mechanical calculator, constructed the empirical sampling distribution, and then approximated it mathematically. Fifteen years later, Fisher grudgingly admitted that he actually got it right. Ever since Gosset, statistics texts and teachers have behaved as if his result was more important than his method. Should we re-think that conclusion? Resampling methods that build empirical sampling distributions the way Gosset did are now easy to use. We'll discuss pros and cons ofresampling and simulation methods in the introductory statistics course and look forward to a discussion of whether the time has come.

Is The T-Test Really Dead? Part 2

✦Paul Velleman, Cornell University, , *pfv2@cornell.edu*; Richard De Veaux, Williams College

Key Words: Gossett, t-test, resampling, introductory course

Student (W. S. Gosset) showed that we only needed the mean, the standard deviation, and the sample size to describe the sampling distribution of the sample mean, with the additional assumption that the population was Normal and a sample size large enough to justify the asymptotic result. Gosset's result is often at the center of our introductory courses. But have we learned the wrong lesson? Gosset derived his sampling distribution by simulating. Using 3000 observations that seemed to be Normal, he drew 750 samples of size 4, computed their means and standard deviations with a mechanical calculator, constructed the empirical sampling distribution, and then approximated it mathematically. Fifteen years later, Fisher grudgingly admitted that he actually got it right. Ever since Gosset, statistics texts and teachers have behaved as if his result was more important than his method. Should we re-think that conclusion? Resampling methods that build empirical sampling distributions the way Gosset did are now easy to use. We'll discuss pros and cons ofresampling and simulation methods in the introductory statistics course and look forward to a discussion of whether the time has come.

Using Technology Resources To Teach Sampling Distributions In Introductory Statistics Courses

✦Ramon Gomez, Florida International University, 813 NW 133rd CT, Miami, FL 33182, *gomezra@fiu.edu*

Key Words: Statistics education, Sampling distributions, SPSS, PowerPoint

The content for a first statistics course at college level typically comprises three logical units: descriptive statistics, probability and inferential statistics. The topic of sampling distributions is presented as a link between probability and the inferential unit. Statistics educators believe that sampling distributions is one of the more difficult concepts to understand by students taking this type of course. In addition to illus-

trating the central limit theorem, the present author's approach while teaching the sampling distributions of means and proportions consists of focusing on the standard errors in such a way that the precision of estimations and significance of hypothesis tests are better understood. This paper discusses the benefits of using technology resources to support this approach with a simulation-based method. SPSS generates the sampling simulations and PowerPoint provides students with the visual comprehension, combining for a highly effective teaching-learning method.

Spatial Statistics And Cartography Using Socr

✦Dave A. Zes, UCLA Department of Statistics, 8125 Math Sciences Bldg., Box 951554, Los Angeles, CA 90095, *davezes@stat.ucla.edu*; Ivo D. Dinov, UCLA Department of Statistics; Nicolas Christou, UCLA Department of Statistics

Key Words: spatial statistics, environmental data, raster maps, vector maps, interpolation, geostatistics

In this paper we present the new spatial statistics and cartography applet of the Statistics Online Computational Resource (SOCR). One of the most appealing aspects of spatial data is their visualization. Using this SOCR applet we have made it easy to import raster and vector maps and overlay data as layers at specific geographic locations defined by their coordinates. Interpolation using the inverse distance method is also available as well as the construction of a raster map using the interpolated values. The analysis of spatial data can be greatly enhanced using maps and geographic information systems tools (GIS). GIS have become popular in recent years and they are encountered in many disciplines. However we have not seen much of their use in statistics, and in particular at introductory level statistics courses. We believe that incorporating spatial data in undergraduate curriculum can engage students and stimulate their interest. Several environmental spatial data sets can be accessed through SOCR.

Helpr: Documentation Made Friendly

✦Barret E Schloerke, Iowa State University Statistics, 180 Brannan St. Apt. 107, San Francisco, CA 94107 US, *schloerke@gmail.com*; Hadley Wickham, Rice University

Key Words: help, R, documentation, html, internet

This talk will describe my 2010 summer research work with Hadley Wickham, on the R package `helpr`. helpr expidites the user finding help in their own working environment. It provides documentation in a friendly HTML format, providing links to other packages, function aliases, and function sources. Finding information about a package and dependencies is painless. There is a also a comprehensive search tool for searching across all R packages for features. Most of `helpr` does not require the internet, while some features (search, RSS feed) require an available internet connection.

Teaching Introductory Statistics Made Easy With R

✦Shiju Zhang, St Cloud State University, 56301, *szhang@stcloudstate.edu*

Key Words: Simulation, Introductory Statistics, R Software

We demonstrate the use of R in teaching introductory statistics courses. Examples include the simulation of sampling distributions, the coverage rate of confidence intervals, and probability.

Basic Statistics With Calculus: Augmenting The Familiar

✦robert jernigan, american university, 4400 massachusetts ave, NW, Gray 202, washington, DC 20016, *jernigan@american.edu*

Key Words: simulation, probability, economics

A bachelor's degree program in mathematics and economics has brought first year students with AP credit into this new course. Students with very basic calculus tools (optimizing a function and finding the area under a polynomial) follow traditional basic statistics topics augmented with calculus tools and online software. For example, sums of absolute and squared deviations are examined in simulations to find the median and the mean, optimal histogram bin widths are found, and least squares estimates are derived. Probability is explored through simulations and calculus of pdfs, moments, and only the use of the moment generating functions (students found this to be the easiest topic). Joint distributions are illustrated with simulated data (double integrals are not used, nor transformations needing jacobians). Independence is examined with simulations and discrete examples. Chebyshev's inequality is proved as well as an outline proof of the central limit theorem. Sampling distributions are examined using simulations. Hypothesis testing of single and pairs of means and proportions are examined. Overall, students put their calculus skills to use in statistics.

525 Methods for Large Spatio-temporal Datasets ●

Section on Statistics and the Environment

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Comparing Methods Of Likelihood-Based Approximate Estimation For High-Dimensional Spatio-Temporal Covariances

✦Yun Bai, Department of Biostatistics, University of Michigan, MI 48109, *yunbai@umich.edu*; Peter Song, University of Michigan

Key Words: high-dimensional, approximate likelihood, spatio-temporal, covariance tapring, composite likelihood

In the recent literature, a number of approximate estimation methods have been proposed to achieve computational feasibility when dealing with high-dimensional covariance matrices for massive spatial and/or spatio-temporal processes. Some popular approaches include covariance tapering, conditional and marginal composite likelihood approaches, and spectral approximations, among others. To date, each of these methods has been only studied in detail in their respective papers, and it remains unknown regarding relative performances of these methods. It is natural to address the understanding of pros and cons of these methods by thorough comparisons. In this paper, based on a comprehensive simulation-based comparison, we report merits and limits of these methods under different settings of spatial and temporal dependence structures.

Non-Parametric Spatial Low Rank Models

✦Cheng Liu, Purdue University, Department of Statistics, 250 N. University Street, West Lafayette, IN 47907-2066, *liu105@purdue.edu*; Hao Zhang, Purdue University

Key Words: EM algorithm, Geostatistics, Kriging, Non-parametric model

We propose a non-parametric approach to modeling the spatial covariance function. The non-parametric approach avoids the selection of parametric family of covariance function. We employ the EM algorithm to estimate the covariance function. Our approach is capable of handling large spatial data sets. Some numerical results are shown to compare our approach to some other approaches.

Reduced Rank Spatial Model With Temporal Confounding Effects

✦Rajib Paul, Western Michigan University, Department of Statistics, 1903 W Michigan Avenue, Everett Tower, Kalamazoo, MI 49008-5278, *rajib.paul@wmich.edu*; Eddy Campbell, The Commonwealth Scientific and Industrial Research Organisation ; John Parslow, The Commonwealth Scientific and Industrial Research Organisation ; Lawrence Murray, The Commonwealth Scientific and Industrial Research Organisation ; Emlyn Jones, The Commonwealth Scientific and Industrial Research Organisation

Key Words: Cross Covariance, Process Convolution, Threshold Estimation

Recent advances in technology facilitate collecting data for ocean processes, such as, salinity, temperature, dissolved oxygen, and chlorophyll. Glider is one of those easy and cheap devices for collecting data. In this paper, we analyze the data collected on the aforementioned processes over Storm Bay in Central Tasmania. The datasets exhibit high spatial resolution, but each location was observed at different time points. Due to advection, these processes can change over a short period of time. Hence, we cannot ignore the time effect completely and use a pure spatial model. We adjust for the time effects by using an autoregressive process of order one in our reduced rank spatial model. Our inferences on the model parameters are based on an MCMC algorithm.

Selection Of Rank And Basis Functions In The Spatial Random Effects Model

✦Jonathan Ray Bradley, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, *bradley.324@osu.edu*; Noel Cressie, Ohio State University; Tao Shi, The Ohio State University

Key Words: Fixed rank kriging, model selection, generalized cross validation, generalized degrees of freedom, spatial statistics, information criteria

Use of the Spatial Random Effects (SRE) model has been motivated by the need to reduce the amount of computation involved in obtaining optimal predictors and their prediction standard errors. This reduction in computation relies on the covariance matrix of the random effects being of low rank. In this paper, the selection of the rank and accompanying basis functions of the SRE model is investigated. We explore selection criteria that include the generalized degrees of freedom and several information criteria; their relative performance is assessed by simulation.

Bayesian Hierarchical Spatio-Temporal Smoothing For Massive Datasets

✦Matthias Katzfuss, University of Heidelberg, Department of Applied Mathematics, Im Neuenheimer Feld 294, Heidelberg, 69120 Germany, *katzfuss@gmail.com*; Noel Cressie, Ohio State University

Key Words: Bayesian Hierarchical Modeling, Remote Sensing, Large Dataset, Dimension Reduction, Varying Model Dimension, Global CO2

Spatio-temporal statistics is prone to the curse of dimensionality: One manifestation of this is inversion of the data-covariance matrix, which is not in general feasible for very-large-to-massive datasets, such as those observed by satellite instruments. This becomes even more of a problem in fully Bayesian models, where the inversion typically has to be carried out many times in an MCMC sampler. We propose a Bayesian hierarchical spatio-temporal random effects (STRE) model that offers fast computation: Dimension reduction is achieved by projecting the process onto a basis-function space of low, fixed dimension, and the temporal evolution is modeled using a dynamical autoregressive model in time. We develop a multiresolutional prior for the propagator matrix that allows for random sparsity and shrinkage. Sampling from the posterior distribution can be achieved in an efficient way, even if this matrix is very large. Finally, we compare inference based on our fully Bayesian STRE model to inference based on EM estimation of the parameters. The comparison is carried out in a simulation study and on a real-world dataset of global satellite CO2 measurements.

Downscaling Temperatures Over The Antarctic Using A Dimension Reduced Space-Time Modeling Approach

✦Jenny Brynjarsdottir, The Ohio State University, 1958 Neil Avenue, 404 Cockins Hall, Columbus, OH 43210 United States, *brynjarsdottir.1@osu.edu*; L Mark Berliner, Ohio State University

Key Words: dimension reduction, statistical downscaling, spatio-temporal modeling

We consider dimension reduced approaches to space-time analysis based on modeling the spatial structure in terms of a low number of specified basis functions. The temporal evolution of the space-time process is then modeled through the amplitudes of the basis functions. We will discuss ways to extend these ideas to modeling two space-time processes and show an example of downscaling temperatures over the Antarctic.

Spatial Anova Modeling of High-Resolution Regional Climate Model Outputs in NARCCAP

✦Emily L Kang, Statistical and Applied Mathematical Sciences Institute, 232 Aaron Circle, Durham, NC 27713, *lei@samsi.info*; Noel Cressie, Ohio State University

Key Words: ANOVA, atmosphere-ocean general circulation model (GCM), North American Regional Climate Change Assessment Program (NARCCAP), regional climate model (RCM), Spatial Random Effects (SRE) model

We consider the current (1971-2000) and the future (2041-2070) average seasonal air surface temperature fields from two regional climate models (RCMs) driven by the same atmosphere-ocean general circulation model (GCM) in North American Regional Climate Change Assessment Program (NARCCAP). We analyze the differences between the future and current temperature fields and incorporate the factor of season and the factor of RCM. Noticing that classical ANOVA doesn't include "spatial" modeling and can't account for the spatial variability across the domain, we propose to use the Spatial Random Effects (SRE) model and carry out ANOVA in the reduced space. Using the SRE model enables us to model the spatial dependence through the spatial basis functions, and our analysis indicates that the interaction between season and RCM is very important. Additionally, with the SRE model, the computation associated with analyzing these high-resolution outputs from RCMs can be carried out efficiently, due to the fixed number of spatial basis functions and the resulting dimension reduction.

526 Recent Advance in Semi-Parametric

Section on Statistics in Epidemiology, Biometrics Section, Section on Nonparametric Statistics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

A Data-Adaptive Approach For Modeling Propensity Scores

✦Yeying Zhu, Pennsylvania State University, Department of Statistics, The Pennsylvania State University, University Park, PA 16803, *yxz165@psu.edu*; Debashis Ghosh, Pennsylvania State University

Key Words: Causal inference, Logistic regression, Random forests, Two-stage modeling, Bias-variance, Observational data

In non-randomized observational studies, estimated differences between treatment groups may arise not only due to the treatment but also because of the masking effect of confounders. To adjust for confounding due to measured covariates, the average treatment effect is often estimated conditioning on propensity scores. In the literature, propensity scores are usually estimated by logistic regression. Alternatively, one can employ non-parametric classication algorithms, such as various tree-based methods or support vector machines. In this talk, we explore the effect of classification algorithms used to model propensity scores using ideas of bias and variance. In addition, we explore ways to combine logistic regression with nonparametric approaches to estimate propensity scores. Simulation studies are used to assess the performance of the newly proposed method and a data analysis example is presented in the end.

Semi-Parametric Estimation For Marginal And Small Data Expression: The Kidney Data Example

✦Norou Diawara, Old Dominion University, 23529, *ndiawara@odu.edu*

Key Words: Bivariate distribution, Dirac delta, likelihood, EM algorithm

The exponential distribution is one of the most used types of distribution because of its importance in many lifetime applications and its properties. So is its bivariate form. Simply used, there can be limitations especially for a heterogeneous type population. A bivariate exponential model with exponential marginal densities, correlated via random variables and with finite probability of simultaneous occurrence. We include the presence of covariate information through a linear relationship, a special case of which was considered by Marshall and Olkin. We extend this model to the bivariate Erlang type distributions with same shape parameter. Illustration with real kidney data is presented. Differences in the kidney disease types are shown in a less tedious method than what has been previously used.

An Estimating Equation For Semiparametric Frailty Models With Error-Prone Covariates

✦Se Hee Kim, Department of Biostatistics, Harvard School of Public Health, 677 Huntington Ave., Boston, MA 02115, *shkim@hsph.harvard.edu*; Yi Li, Dana-Farber Cancer Institute and Harvard School of Public Health ; Donna Spiegelman, Harvard School of Public Health

Key Words: copula, estimating equation, measurement error, proportional hazards model

We consider semiparametric frailty models for the correction of measurement errors in survival data analysis. An estimating equation method based on the conditional expectation of the unobservable frailty is proposed for regression coefficients in the proportional hazard model with error-prone covariates. We allow general structures for the error-prone covariates, not restricted to a linear additive measurement error model or a Gaussian measurement error. The conditional distribution of the frailty given observed covariates is estimated nonparametrically through a copula as a function of the marginal distributions of the true exposure and the surrogate. The proposed approach via the copula can enhance the precision of estimation by utilizing all the available information on the surrogate from both main and validation studies. Asymptotic properties are established well, and finite sample properties are evaluated through extensive simulation studies. The proposed method is applied to data from the Nurses' Health Study and Health Professionals Follow-up Study.

Generalized Multiple Indicators, Multiple Causes Measurement Error Models

✦Carmen Dwele Tekwe, Texas A&M University, Department of Statistics, College Station, TX 77840, *ctekwe@stat.tamu.edu*; Randy L Carter, University at Buffalo

Key Words: Berkson error, measurement error, MIMIC models, nonlinear models, structural equation models, atomic bomb survivors

Multiple indicators, multiple causes (MIMIC) models are useful for studying the effects of a latent variable on several outcomes, when causes of the explanatory latent variable are observed. Classical measurement error is uncorrelated with the latent variable; while a Berkson error is uncorrelated with its estimate. Previous work has focused on MIMIC models, where the causes of the latent variable are observed without error. We generalize the MIMIC model to allow non-linear relationships and also allow both Berkson error and classical measurement error in the determinants of the causal equation by defining the

G-MIMIC ME model. We propose MC-EM based estimation procedures for the G-MIMIC ME models and apply our results to data collected on atomic bomb survivors.

Simultaneous Confidence Bands For Additive Models With Locally Adaptive Smoothed Components And Heteroscedastic Errors

✦Manuel Wiesenfarth, Georg-August-Universitaet Goettingen, Wilhelm-Weber-Str. 2, Goettingen, International 37073 Germany, *mwiesen@gwdg.de*; Tatyana Krivobokova, Georg-August-Universitaet Goettingen; Stephan Klasen, Georg-August-Universitaet Goettingen

Key Words: confidence band, penalized splines, volume-of-tube formula, adaptive estimation, mixed model, nonparametric regression

We propose a simple and fast approach to construct simultaneous confidence bands for smooth curves that enter an additive model, are spatially heterogeneous and are estimated from heteroscedastic data. Estimation is based on the mixed model representation of penalized splines which allows to fit such complex models from the corresponding likelihood and helps to build simultaneous confidence bands using the approximation to the tail probability of maxima of Gaussian processes. These confidence bands have very good small sample properties and are obtained instantly, i.e. without using bootstrap. Based on the resulted confidence bands a lack-of-fit test is proposed which not only performs competitively compared to likelihood ratio tests, but also allows to incorporate the above mentioned model features without any additional effort. Finite sample properties are studied in simulations and an application to undernutrition in Kenya shows the practical relevance of the approach. The method is implemented in the R package AdaptFitOS.

Efficient Targeted Estimation Using Instrumental Variables

✦Boriska Toth, UC Berkeley, 1540 Milvia St. #6, Berkeley, CA 94709 United States, *boriska.toth@gmail.com*; Mark van der Laan, University of California, Berkeley

Key Words: semiparametric estimation, causal inference, clinical trials, instrumental variables, targeted maximum likelihood, treatment effect

The problem of estimating the causal effect of a variable on an outcome when error terms are dependent on the covariates is of fundamental importance. The presence of valid instrumental variables allows for unbiased estimates of the causal effect. To our knowledge, this work is the first to empirically study the use of targeted maximum likelihood-based estimation (TMLE) to estimate a causal effect using an instrument. TMLE is an approach for unbiased, asymptotically efficient semiparametric estimation, with additional desirable properties. We demonstrate the construction of TMLE estimators for a general setting involving instruments (which captures many common scenarios in which one wishes to estimate the effect of treatments). In simulation data having unmeasured confounding, we compare empirically the performance tradeoffs between a model using instruments vs a model invalidly based on conditional independence assumptions. We further compare the TMLE-based estimators to the widely-used two-stage regression-based approaches. Finally, we perform analysis on real-life biostatistics data using the important example of estimating the effect of treatment under non-compliance.

527 Contributed Oral Poster Presentations: Section on Nonparametric Statistics

Section on Nonparametric Statistics

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Asymptotic Normality for the Kernel Estimator of the Regression Function for Censored Time Series

✦Zohra Guessoum, USTHB, universite des sciences et de la technologie Houari Boumediene, FacultÈ de MathÈmatiques, dÈpt probabilitÈs, statistiques,BP 32, El Alia, Bab Ezzouar, Algiers, 16111 Algeria, *zguessoum@usthb.dz*; Elias Ould Said, Lab. LMPA Univ. Lille Nord de France

Key Words: censored data, asymptotic normality, strong mixing, nonparametric regression

In this paper, we consider the estimation of the regression function when the interest variable is subject to random censorship and the data satisfy some dependency conditions. We show that the kernel estimate suitably normalized is asymptotically normally distributed and the asymptotic variance is given explicitly. An application to confidence bands is given. Some simulations are drawn to lend further support to our theoretical results and to compare finite samples sizes with different rates of censoring and dependence.

Improved Nonparametric Mean And Variance Estimators For Judgment Post-Stratification

✦Jesse Frey, Villanova University, 800 Lancaster Ave, Villanova, PA 19085, *jesse.frey@villanova.edu*; Timothy Feeman, Villanova University

Key Words: Judgment post-stratification, Ranked-set sampling, Admissibility, Variance estimation

We consider nonparametric estimation of the population mean and population variance using judgment post-stratification (JPS). We prove that the standard nonparametric JPS mean estimator is inadmissible under squared error loss within a certain class of linear estimators. We then derive alternate mean estimators that are admissible in this class, and we show that one of them is always better than the standard estimator. We also derive a new nonparametric JPS variance estimator that seems to outperform the JPS analogue of a well-known nonparametric variance estimator for ranked-set sampling.

Efficient Rank Regression With Wavelet Estimated Scores

✦Eddy Armand Kwessi, Auburn University, 549 East Glenn Avenue, #15, Auburn, AL 36830 USA, *kae0005@tigermail.auburn.edu*

Key Words: Efficient, Rank, Regression, Wavelet, Scores

We will discuss asymptotically efficient rank estimation of linear models. The approach is based on the estimation of the score function in the rank dispersion function using compactly supported wavelets. As a by-product of this approach, we obtain a consistent estimator of the asymptotic variance of the rank estimator. We will discuss extensions of this result to mixed effect models

An Empirical Study Of A Rank-Based Estimate For Cell Lineage Data

✦Tamer Elbayoumi, Department of Statistics, 3304 Everett Tower, Western Michigan University, Kalamazoo, MI 49008, *tamer.m.elbayoumi@wmich.edu*; Jeffrey Terpstra, Department of Statistics

Key Words: Rank-Based Estimates, Least Squares Estimate, Bifurcating Autoregressive, Cell Lineage Data, Sample Relative Efficiency, Aberrant Observations

The presence of aberrant observations (i.e. outliers) in cell lineage data is quite common. As such, it is desirable to have an outlier-resistant estimation procedure as an alternative to least squares estimation (maximum likelihood estimation under normality). In this work, we consider rank-based estimates of the parameters of a first order bifurcating autoregressive [BAR(1)] model. The BAR(1) model was proposed by Cowan and Staudte (1986) for cell lineage data. In it, each line of descendents follows a first order autoregressive [AR(1)] model and allows sister cells from the same mother to be correlated. Real examples and a simulation study are performed in order to examine the behavior of these rank-based estimation procedures. More specifically, we compute finite sample relative efficiencies with respect to least squares estimate. The results indicate that the rank-based estimation procedures are more efficient when outlying observations are present.

Rank Based Variable Selection For Classification

✦Shuxin Yin, Auburn University, 221 Parker Hall, Mathematics and Statistics, Auburn University, AL 36849 United States, *yinshux@auburn.edu*; Asheber Abebe, Auburn University

Key Words: Wilcoxon-Mann-Whitney, Principal components, Gene selection, Misclassification error, Transvariation

Gene expression microarray data are often used to classify samples into tumor categories. It is known that standard statistical methods in classification are inefficient particularly when sample size is larger than the data dimension. Gene expression data are often ultra-high dimensional. This calls for methods of selecting the genes or gene combinations that are the most informative for class prediction in the sense that the resulting genes or gene combinations result in small misclassification error rates. We propose a novel variable selection method using the Wilcoxon-Mann-Whitney type statistics to measure group separation and overlap. Misclassification error rates of two gene expression data sets are found using six different classifiers. Prior to classification, genes were selected by using the the proposed Wilcoxon-Mann-Whitney selection method and gene components that were picked by PCA and PLS. A Monte Carlo simulation study demonstrates that the proposed method performs better than PCA and PLS for heavy-tailed distributions whereas it does worse than PCA and PLS for normally distributed data.

Recursively Imputed Survival Trees

✦Ruoqing Zhu, UNC-Chapel Hill, 3105 McGavran-Greenberg Hall, CB #7420, Chapel Hill, NC 27599-7420, *rzhu@bios.unc.edu*; Michael R Kosorok, UNC-CH

Key Words: Trees, Ensemble, Survival Analysis, Variable Selection, Imputation

Tree-based methods have become increasingly popular statistical tools since Breiman et al. (1984) introduced the classification and regression tree (CART) to the statistics research community. We adept tree-based method into survival arena and yield a highly accurate nonparametric survival model for right censored survival data: Recursively Imputed Survival Tree (RIST) regression. A novel censoring imputation method is purposed to generate failure times for censored observations. This imputation method utilize censored observations in a way like no previous method and result in a better model fitting and prediction. We also discuss the consistency of tree-based method and its future application in variable selection.

Follow-Up Testing In Functional Anova

✦Olga A. Vsevolozhskaya, Montana State University, 1836 E. Valley Center Rd., Apt A, Bozeman, MT 59718, *tenar@live.com*; Mark Greenwood, Montana State University

Key Words: Functional data analysis, Permutation methods, Multiple comparison procedures

Functional analysis of variance involves testing for differences in functional means across k groups in n functional responses. If a significant overall difference in the mean curves is detected, one typically would like to identify the regions of difference more specifically. Two different follow-up testing methods are discussed and contrasted. A previous point-wise permutation adjustment method is compared to a new method based on assessing significant differences within regions of the functional responses. The methods are contrasted both in a simulation context and with an application.

A Non-Least-Squares Method For Estimation Of Broken-Line Regression To Determine Plateau Time Of Outputs In Some Animal Science Studies

✦Greg Knofczynski, Armstrong Atlantic State University, Department of Mathematics, 11935 Abercorn St, Savannah, GA 31419, *greg.knofczynski@armstrong.edu*; Lorrie Hoffman, Armstrong Atlantic State University

Key Words: change point, joinpoint, nonparametric

Broken-line regression has been used to estimate parameters in increasing or decreasing linear trends that later plateau. Most approaches employ least squares methods that under usual assumptions lead to maximum likelihood estimates (Gill, 2004). In Hoffman, Knofczynski and Clark (2010) an estimation procedure called MMNPR is derived and simulation studies show that parameter estimates of the broken-line regression perform better than the MLEs in small sample situations when data is from distributions where error terms are ill-behaved. A review of applied science literature revealed many studies using broken-line regression. We note three animal studies that report parameters based on least squares approaches: 1) determining times of steady

milk production in cows (Sahinler, 2009), 2) transport of urea in mice (Marini, Lee and Garlick, 2006) and 3) the maximum sustained effect of dietary copper on chicks (Robbins, Norton and Baker, 1979). We investigate the appropriateness of underlying assumptions concerning correlation and homoscedacity and revisit the analysis of the animal data using the MMPNR technique, discussing differences.

Tests For Stationarity Based On Minors And Singular Values Of The Nonstationary Spectrum

✦Azadeh Moghtaderi, Queen's University, University Avenue, Jeffery Hall, Queen's University, Kingston, ON K7L3N6 Canada, *azadeh.moghtaderi@gmail.com*; Glen Takahara, Queen's University

Key Words: Stationary, Nonstationary, Statistical Test, Spectrum, Time-Frequency, Rank

We consider the Evolutionary Spectrum (ES) of a nonstationary process, which in general is a function of both time and frequency. For stationary processes the ES is time independent and for Uniformly Modulated Processes (UMPs, i.e., a stationary process multiplied by a positive deterministic function of time, the modulating function) the ES as a function of time is the same function up to a proportionality constant, which depends on the frequency under consideration. By considering the ES on any 2-dimensional finite grid of time-frequency values, we obtain a matrix which has rank 1 in both cases and rank larger than 1 in general otherwise. Moreover, taking the logarithm of any such matrix gives a log-spectrum matrix which also has rank 1 in the stationary case but rank 2 in the UMP case, and rank greater than 2 otherwise. Based on these observations, we develop graphical and statistical tests based on the eigenvalues or the 2 by 2 minors of an estimator of the spectral and log-spectral matrices. We illustrate the performance of our testing procedures via simulation examples and provide some distributional results for the test statistics we consider.

Robust Principal Component Estimation Of Functional Logistic Regression

✦Melody Denhere, Auburn University, Department of Mathematics and Statistics, 221 Parker Hall, Auburn, AL 36849, *mbd0002@tigermail.auburn.edu*; Nedret Billor, Auburn University

Key Words: functional data analysis, robust statistics, functional logistic regression, principal components

We propose a robust estimator for the parameter function in the functional logistic regression model by using robust principal component approach. Robustness properties of the proposed robust estimator are explored via the robustness measures. A numerical implementation is also given with the performance of the proposed robust estimator evaluated both in a simulation study and with real datasets.

Improved Critical Values For The Two Independent Samples Winsorized T Test

✦Shlomo Sawilowsky, Wayne State University, Detroit, MI 48202, *snbaay@yahoo.com*; Piper A Farrell-Singleton, Wayne State University

Key Words: Winsorized t, Monte Carlo, Critical Values

Since 1963, critical values for conducting the two independent samples Winsorized t test have been based on adjustments to the df in order to obtain estimated critical values from the t table by Tukey and McLaughlin. The purpose of this poster is to present improved critical values obtained from a Monte Carlo program written in Visual Basic, with 1,000,000 repetitions per critical value, based on pseudo-random numbers drawn from the normal distribution. Thirty Apple computers with dual core Intel 2.66 GHz CPUs were employed to conduct the simulation. Estimated critical values for both one and two tailed tests were identified at the 95th, 97.5th, 99th and 99.5th percentiles. Critical values were tabled for sample sizes per group of n=5, 6,...30, 45, 60, 90, and 120, specifically for df=8,10,...18 (symmetric Winsorizing (W) of W=1 per side; 20, 22,...26 (W=1, 2); 28, 30,...36 (W=1, 2, 3); 38, 40,...46 (W= 1,2.4); 48, 50,...56 (W=1, 2,.5); 58 (W=1,2,.6); 88 (W=1,2,.9); 118 (W=1,2,.12); 178 (W=1,2,.18); and 238 (W=1,2,.24).

Type I And Ii Errors For Improved Critical Values For The Winsorized T-Test

✦Michael Lance, 15633 Northville Forest Drive, Apt. Q182, Plymouth, MI 48170, *michael.lance@gmail.com*; Shlomo Sawilowsky, Wayne State University

Key Words: Winsorized t, Monte Carlo, Critical Values

Since Tukey and McLaughlin (1963), it has been accepted practice for critical values for the Winsorized t test for two independent samples to be based on adjusted degrees of freedom from the t table. Recently, a new table of Winsorized critical values was developed via Monte Carlo methods. Based on five mathematical distributions (Cauchy, Chi-squared (df=2), exponential, t(df=3), uniform), and eight common data sets from social and behavioral sciences (Micceri, 1989), the different critical values were compared with respect to robustness to Types I and II errors, using Monte Carlo methods via a Fortran program with 1,000,000 iterations per experiment. The improved critical values yielded better Type I and II error results for small samples than did Tukey and McLaughlin's critical values, until the sample sizes reach a magnitude of about n=90 per group, at which point the two tables of critical values yielded similar results.

Rank-Based Approach For Mixed Effects Hierarchical Models In Nested Designs

✦Yusuf K Bilgic, Western Michigan University, 4530 lilac ln. Apt.42, Kalamazoo, MI 49006, *yusuf.k.bilgic@wmich.edu*

Key Words: Rank-based analysis, Mixed Hierarchical Model, Nested Design, Educational Research

Hierarchical designs frequently occur in many areas. The experimental design of interest is expressed in terms of fixed effects but, for these designs, nested factors are a natural part of the experiment. These nested effects are generally considered random and must be taken into account in the statistical analysis. In many of the applications that we have considered outliers frequently occur. Traditional analyses are quite sensitive to these outliers and lose considerable power to detect the fixed effects of interest. We offer two new rank-based algorithms, Rank Prediction Procedure (RPP) and Iterative Generalized Rank Estimation Procedure (IGREP), which iteratively obtain robust estimation for both the fixed and random effects. We present the results of a small sample investigation of these analyses, including a comparison with traditional analyses.

Simultaneous Confidence Intervals For Location Shift Parameters Using Permutation Tests

✦Scott Richter, University of North Carolina at Greensboro, Greensboro, NC 27412, *sjricht2@uncg.edu*; Melinda McCann, Oklahoma State University

Key Words: Location shift, simultaneous confidence interval, permutation test

A method for computing simultaneous confidence intervals for location shift, for all pairwise comparisons, is presented. The reference distribution is based on the permutation distribution of the maximum absolute pairwise difference from among all pairs. The method guarantees strong control of the familywise confidence level, and does not require assumptions about the form of the population distribution.

Nonparametric Regression Modeling Of Extreme Events

✦Bet¸l Kan, Anadolu University, Yunus Emre Kamp¸s¸, Fen Fak¸ltesi, Istatistik Bˆl¸m¸, Eskisehir, 26470 Turkey, *bkan@anadolu.edu.tr*; Ahmet Sezer, Anadolu University; Berna Yazici, Anadolu University

Key Words: Splines, Likelihood Ratio Test, nonparametric Regression, Extreme Events

Extreme value data analysis has been widely used in nonparametric regression modeling. This study aims to introduce the fundamentals of extreme value theory as well as practical aspects for estimating and evaluating the models for tail related risk measures. The nonparametric and the linear models are taken into account to form the relationship between the temperature and time effect. The log-likelihood value has been used to evaluate the adequacy of the form of the covariate effects on real life data set. For each year from 1991 to 2010 the 15 highest daily temperature values are analyzed by using the block Gumbel distribution to investigate the global warming.

Constrained Penalized Splines Regression In The Presence Of Correlated Errors

✦Huan Wang, Colorado State University, , *wangh@stat.colostate.edu*; Mary C Meyer, Colorado State University; Jean Opsomer, Colorado State University

Key Words: shape constraints, penalized splines, AIC, correlated errors

We propose using penalized spines with shape constraints such as monotonicity or convexity for estimation and inference of regression functions in presence of correlated errors. Error structures include AR(p), MA(q), and ARMA(p,q). The parameters p and q as well as the penalty parameter are chosen by the AIC criterion. Simulations for comparing the behaviors of AIC, estimation of error structure and fit of the data for constrained penalized splines fit, unconstrained penalized splines fit and simple linear regression fit are conducted. The results show that our proposed estimation method behaves better than the other two methods in both estimating the structure of errors and mean function of the trend.

Thyroid Hormone Treatment Efficacy In Bipolar Disorder With Irregular And Sparse Cyclical Longitudinal Data

✦Brian Calimlim, University of California, Los Angeles, 5573 Olympic Blvd., Los Angeles, CA 90036, *brian007@gmail.com*; Catherine Ann Sugar, University of California, Los Angeles; Patricia Walshaw, University of California, Los Angeles

Key Words: Bipolar Disorder, Markov Chains, Permutation Tests, Cyclicity, Longitudinal Data, Depression/Mania

Bipolar Disorder (BD) is a psychiatric illness in which a person experiences episodes of elevated mood and energy, known as mania, and episodes of lowered mood and interest, known as depression. Due to its complex nature, challenges arise in analyzing BD data. Irregular cyclical patterns, varying in amplitude and phase by subject, make it difficult to develop simple parametric models, while the dual nature of BD requires a multi-component outcome measure. Further complications exist in treatment trials as efficacy can be manifested in many ways. In a study of thyroid hormone treatment efficacy for refractory rapid cycling BD, 32 patients were randomized to receive triiodothyronine (T3), thyroxine (T4), or a placebo plus usual medication. Depressive and manic symptoms data were longitudinally gathered before and after treatment. We present an analysis of this study and strategies taken to handle data irregularity and sparsity. A Markov chain approach is used for modeling, while permutation tests are developed to determine statistical significance. Preliminary results suggest patients receiving T3 show reductions in mania while those on T4 improve with regard to depression.

A Nonparametric Approach To Spatial-Temporal Functional Data

✦Nuen Tsang Yang, University of California, Davis, 1 Shields Avenue,, Davis, CA 95616, *matyang@ucdavis.edu*

Key Words: functional, geostatistis, nonparametric, spatial-temporal, fMRI

Functional data analysis with spatial dependence required the tools of spatial statistics and functional data analysis. Multivariate spatial statistical tools are generalized to the case of functional data. In spatial functional analysis, one of the interests is to use functional data to model spatial dependence. In this paper, a new high dimensional nonparametric method is introduced to model the covariance structure of spatial-temporal functional data. The notion is to generalize the variogram function to nonparametric form under some regularity assumptions. Different smoothing methods are used and evaluated by nonparametric and functional cross-validations. For illustration, the method is applied to California temperature data and brain image data from the fMRI Data Center.

528 Contributed Oral Poster Presentations: Section on Statistical Computing

Section on Statistical Computing

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Asymptotic Expansions for the Pivots Using Log-Likelihood Derivatives with an Application in Item Response Theory

✦Haruhiko Ogasawara, Otaru University of Commerce, Otaru, International 047-8501 Japan, *hogasa@res.otaru-uc.ac.jp*

Key Words: pivots, log-likelihood derivatives, inverse expansion, sandwich estimator, item response theory

Asymptotic expansions of the distributions of the pivotal statistics involving log-likelihood derivatives under possible model misspecification are derived using the asymptotic cumulants up to the fourth order and the higher-order asymptotic variance. The pivots dealt with are the studentized ones by the estimated expected information, the negative Hessian matrix, the sum of products of gradient vectors, and the so-called sandwich estimator. It is shown that the first three asymptotic cumulants are the same over the pivots under correct model specification with a general condition of the equalities. An application is given in item response theory, where the observed information is usually used rather than the estimated expected one.

Parameter Estimation And Group Classification With Mixture Item Response Theory Models

✦Holmes Finch, Ball State University, Department of Educational Psychology, Ball State University, Muncie, IN 47306, *whfinch@bsu.edu*; Brian French, Washington State University

Key Words: mixture models, item response theory, bayesian modeling, maximum likelihood, group classification

Mixture item response theory models are increasingly popular in psychometrics for identifying (a) differentially functioning items for various item types and (b) latent groups. The two most popular methods for estimating these models are maximum likelihood (MLE) and Markov Chain Monte Carlo Bayesian techniques (MCMC). Although popular, no systematic investigation has examined these methods in terms of the accuracy of grouping individuals or estimates of item difficulty and discrimination parameters for groups. This simulation study compared MLE with MCMC in terms of parameter estimation accuracy as measured by root mean squared error (RMSE) and grouping accuracy for varying sample sizes, number of latent classes, number of items, and degree of group separation. Results reveal that both methods correctly classified more than 90% of cases, with MCMC being slightly more accurate. RMSE for item difficulty estimation was comparable for the two methods. MLE produced lower RMSE values, indicating more accurate estimates, for item discrimination. The complete paper will discuss these results in detail and offer recommendations for practitioners applying these methods.

A Moment Preserving Finitization Across The Power Series Family Of Probability Distributions

✦Martin Stuart Levy, University of Cincinnati, PO Box 210130, Department of Operations and Business Analytics, Cincinnati, OH 45221, *martin.levy@uc.edu*; James J. Cochran, Louisiana Tech University; Saeed Golnabi, Optimum Office Solutions

Key Words: Poisson distribution, Taylor series expansion, power series family of distributions, moments, finitization, variate simulation

Finitization transforms a discrete distribution into a distribution with smaller support of specified size. In special cases finitization preserves moments; moments of the order n finitization coincide with those of the first n moments of the parent distribution. We create a moment preserving finitization method for power series distributions by introducing an alternative representation and showing how to finitize members of this new class in a manner that preserves moments of the parent distribution. We provide results on convolutions and a reproductive property for power series distributions that have been finitized in this manner. We describe how these finitized distributions accelerate variate generation in simulation trials.

A Copula Approach To Fit A Regression Model To Interval-Valued Variables

✦Alisson de Oliveira Silva, Federal University of Paraiba, Cidade Universitaria, s/n, Jo„o Pessoa, International 58051-900 Brazil, *allysson_jlr@yahoo.com.br*; Eufr·sio de Andrade Lima Neto, Federal University of Paraiba; Ulisses Umbelino dos Anjos, Federal University of Paraiba

Key Words: Copula, Regression, Interval variable, Symbolic Data Analysis

Symbolic Data Analysis has been introduced as a domain related to multivariate analysis, pattern recognition and artificial intelligence. Interval-valued variables have been mainly studied in the area of SDA. Recently, different approaches have been proposed to fit a regression model for interval-valued variables. Based on the copula theory (Sklar, 1953), we consider a general bivariate distribution for the response variable Y = [Y1, Y2], which can be represented by the lower/upper bounds of the interval or by any pair of interval features. Some of the most important copulas (Gaussian and some Archimedean families) will be considered to modeling the dependence structure between Y1 and Y2. Gaussian, Gamma and Inverse Gaussian are possible candidates for the marginal distributions. The coefficients of the new regression method (CopRIn) will be estimated through BFGS method. The CopRIn method will be applied to a two real interval data sets and your prediction performance will be compared with the methods CRM (Lima Neto and De Carvalho, 2008), CM (Billard and Diday, 2000) and BSRM (Lima Neto et. al, 2010). The new approach outperforms the previous ones based on root mean square error.

Robust Em Clustering Through A Mixture In Multivariate Regression

✦Ji Young Kim, Mount Holyoke College, 50 College Street, South Hadley, MA 01075, *jiykim@mtholyoke.edu*; Xuming He, University of Illinois at Urbana-Champaign; John Marden, University of Illinis at Urbana-Champaign

Key Words: Robust, Clustering, Multivariate Regression, EM akgirutgn

Among a variety of clustering techniques, the mixture model approach provides a sound and flexible option. The normal mixture models fitted through the EM algorithm have been commonly used for clustering. To account for heavier tails, the mixture of t distributions has been used as an alternative. However, data with non-Gaussian tails, asymmetry, or extreme outliers may not be adequately modeled by a mixture of t distributions. In some problems, clustering can be done effectively once after adjusting for certain covariates. We propose a robust clustering method (REM) for high-dimensional data based on the idea of mixtures in a multivariate linear regression setting, designed to reduce the effects of the outliers with Huber's loss function. The model is fitted via the EM (expectation and maximization) algorithm with a new iterative algorithm to solve the optimization problem. In simulation studies, we find that the REM often works better than existing clustering methods. The superior performance of the REM can be attributed to the robustness and the use of covariates. Along with other existing methods, the REM provides an elaborate approach to clustering heterogene

Simulated Likelihood Estimation Of Measurement Error Models

✦Fassil Nebebe, Concordia University, Department of Decision Sciences & MIS (MB 12.119), 1455 De Maisonneuve Blvd. West, Montreal, QC H4V 1Y9 Canada, *fnebebe@jmsb.concordia.ca*

Key Words: importance sampling, incidental parameters, measurement error model, polynomial functions, simulated likelihood

We consider in this paper a unified approach for the estimation of a general measurement error model. The model may contain incidental parameters, though they are assumed to be realizations of latent variables. In this sense, the model considered is "structural", but it differs from the traditional structural relationship in the uses of conditional analyses in the treatment of asymptotic inferences. Since the maximum likelihood function does not have a closed form, numerical methods for obtaining the maximum likelihood estimates are considered. In particular, the method of simulated likelihood based on importance sampling is studied in detail. It is seen that the procedure can be greatly facilitated with an automated choice of an importance function. Asymptotic properties of the simulated maximum likelihood estimates are investigated. The efficiency of the simulated likelihood approach is then compared with those of the traditional methods in the special case of polynomial functional relationships.

The Planet Jupiter- A Case Study

✦Morteza Marzjarani, Saginaw Valley State University, 7400 Bay Road, University Center, MI 48710, *marzjara@svsu.edu*; Morteza Marzjarani, Saginaw Valley State University

Key Words: Planet, Statistical analysis

The study of our planets is always fascinating. The government agency NASA is the leading institution in space exploration. There are still numerous unanswered questions about our planets and it seems there is no end to know all about these planets. Among the planets, Jupiter (the largest gas planet) is of special interest and it is the focus of this study. There are still many unknowns about the structure of this planet. In this study, we would like to present an overview of this planet and analyze data collected by satellites and ground-based devices using some statistical models and interpret the results.

Regression Model Stochastic Search Via Local Orthogonalization

✦Ruoxi Xu, Department of Statistics, The Ohio State Univeristy, 1958 Neil Avenue, 231 Cockins Hall, Columbus, OH 43210, *xu.232@osu.edu*; Chris Hans, Department of Statistics, The Ohio State Univeristy

Key Words: Model Uncertainty, Gibbs, Metropolis Hastings, Othonomral Rotation, Multicollinearity, Point-Mass Prior

Bayesian model uncertainty problems are often challenged by high dimensional model spaces. When the number of predictors is large, it is often infeasible to enumerate the model space, which makes Markov chain Monte Carlo (MCMC) methods such as the Gibbs sampler popular among practitioners. A common problem with the Gibbs sampler is its potential to get stuck in local regions of the model space when predictors are highly correlated. Motivated by the need to explore the model space efficiently when high multicollinearity presents, we introduce a Metropolis-Hastings-Based algorithm with an orthonormal rotation on the regression coefficients. The orthonormal rotation is based on the spectrum decomposition of the covariance matrix and is shown to facilitate MCMC updates in directions along which the chain is less likely to stay in local regions. We demonstrate the effectiveness of the resulting sampling algorithm over other popular sampling methods on a real data set with a large number of highly correlated predictors.

Hierarchical Car Models For Survival Data With Cure Rate Models

✦Yajun Liu, University of Missouri, 134 Middlebus Hall, University of Missouri, Columbia, MO 65211, *ylq89@mail.missouri.edu*

Key Words: CAR, Cure Rate Model, Weibull, Survival, Ratio-of-Uniforms, ARM

In this paper, we proposed a Bayesian hierarchical linear mixed model with spatial effects for right-censored survival data analysis. The model is extended from a special case of the classic Cox model with Weibull hazard and cure rate model. We applied the generalized linear mixed model with different effects (age, county, cancer primary site, cancer situation, etc.) to the hazard rates. A conditional autoregressive (CAR) prior is used to capture the spatial effects, which are decided by counties the survival subjects belong to. The computation is done by Gibbs Sampling. The Ratio-of-Uniforms method is used to sample from a non-standard conditional posterior density, and the ARM is also used to sample from log-concave densities. Sensitivity of the shape parameters in the Weibull distribution for the modeling is also analyzed. DICs are compared between different models. The simulation is done based on the the Colon \& Rectum Cancer (CRC) incidences in Iowa from SEER Data. The computation is implemented in Intel Fortran on Linux platform.

Variable Selection In A Bayesian Additive Trees Model

✦Yue Zhao, University of Massachusetts Amherst, Department of Mathematics and Statistics, University of Massachusetts Amherst, Amherst, MA 01003, *yzhao@math.umass.edu*; John Staudenmayer, University of Massachusetts, Amherst

Key Words: Bayesian additive trees, CART, MCMC, Variable selection

We consider the problem of multiple non-parametric regression when the number of possible covariates is possibly larger than the sample size. We address the problem with an additive trees model where variable selection is a build-in feature. While trees are powerful tools for modeling complex variable dependence, it is difficult to add variable selection to traditional additive trees models due to the nature of unregulated tree growing algorithms. We develop a Bayesian approach that includes a novel prior for the tree that restricts tree size and the size of the search space. This enables us to easily extend the standard Bayesian variable selection technique to the additive trees model. An MCMC sampler is used to generate samples from the posterior. The method is evaluated and compared to other approaches using simulated data.

Simulation Study To Detect The Existence Of Spatial Patterns: An Application Of Randomization Test

◆Denise Nunes Viola, UFBA, Salvador, 40150080 Brasil, *viola@ufba.br*

Key Words: Randomization test, Spatial pattern, Semivariogram, Mantel test

Often the researcher's interest is to model the spatial distribution in different areas. In Geostatistics, semivariogram is commonly used to verify the existence of spatial patterns, but eventually it should not be used. An alternative to identify the existence of spatial patterns when the assumptions are not met is to use an adaptation of the Mantel Randomization Test. The randomization test is appropriate for small samples and can be used for non random samples. This test compares the value of an observed statistic for the original data with the values of the statistic after randomizing the observations, where the p-value is given by the proportion of times the randomized data was statistically greater than or equal to statistics obtained from the original data. To test it, was generated a thousand sets of 250 observations. Data was generated with spatial dependence of 7 meters of range in a 20x20 meters area. The matrix of distances from the location of collection was randomized 10,000 times to get the p-value. In 81.20% of the randomizations it was confirmed the existence of spatial pattern. Additional studies will be conducted to confirm the results.

Condition Number Regularized Convariance Estimation

◆Joong-Ho Won, VAPAHCS, , *JoongHoJohann.Won@va.gov*

Key Words: covariance matrix, convex optimization, regularization, condition number

Estimation of high-dimensional covariance matrices is known to be a difficult problem, has many applications, and is of current interest to the larger statistics community. Several approaches to regularize high-dimensional covariance estimates have been proposed in the "large p small n" setting. In many applications, the estimate of the covariance matrix is required to be not only invertible, but also well-conditioned. Although many regularization schemes attempt to do this, none of them address the ill-conditioning problem directly. In this paper, we propose a maximum likelihood approach, with an explicit constraint on the condition number, with the direct goal of obtaining a well-conditioned estimator. No sparsity assumption on either the covariance matrix or its inverse are are imposed, thus making our procedure more widely applicable. We demonstrate that the proposed regularization scheme is computationally efficient, yields a type of Steinian shrinkage estimator, and has a natural Bayesian interpretation. We investigate the theoretical properties of the regularized covariance estimator comprehensively, including its regularization path.

The Application Of Multivariate Statistics To The Face Recognition Problem

◆Alex Jerald Sawyer, Alabama A&M University, 115 Chase Rd. Apt. 7202, Huntsville, AL 35811, *ajexsawyer@sbcglobal.net*

Key Words: Principle Component Analysis, Linear Discriminant Analysis, Small Sample Problem, Face Recognition

The demand for the use of digital images for face recognition is on the increase. This poster briefly discusses the theories for Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) and how they are combined to perform digital face recognition. The coverage includes a definition of the small sample problem and how that problem is solved through the use of the singular value decomposition procedure. The poster uses an existing database of statistical images of human faces to compare the performances of the PCA and LDA combinations. The database contains 4160 images with variations over lighting, pose, distance and infrared imagery. Face reconstruction quality and face recognition accuracy are evaluated over several variable settings, including sample size.

Generalized Fiducial Inference For An Important Class Of Discrete Distributions

◆Todd Iverson, Saint Mary's University of Minnesota, 700 Terrace Heights, Winona, MN 55987, *tiverson@smumn.edu*

Key Words: fiducial, generalized inference, discrete distributions, regression, generalized linear models

We present a generalized fiducial framework for a class of univariate discrete distributions. This class includes poisson and logistic regression from generalized linear models, but also allows the means for poisson and bernoulli random variables be modeled as a linear function of the explanatory variables. A Monte Carlo method for generating the fiducial distribution for the parameters is illustrated using poisson regression and the results of ongoing simulations comparing this method to standard methods will be summarized.

Application Of Smoothly Truncated Levy Distributions To The Eeg-Sleep Study Of Neonates

◆Alexandra Piryatinska, San Francisco State University, 1000 Continentals way, apt 214, Belmont, CA 94002, *alpiryat@sfsu.edu*

Key Words: levy processes, tempered stable processes, numerical MLE, EEG-sleep data

Smoothly Truncated Levy Flights were introduced by Koponen (95). Rosinski (2007) studied the more general class of processes of similar type which he called tempered stable processes. A tempered stable process combines both the α-stable and Gaussian trends. At a small time scale it is close to an α-stable process while at a large time scale it approximates the Brownian motion. Tempered stable processes found applications in finance and biology. In our study we consider symmetric Smoothly Truncated Levy (STL) distributions and processes. This class of distributions is described via characteristic function and there are no closed form for the density functions and we developed numerical Maximum Likelihood Estimates (MLE) for them. The above method was applied to a study of EEG-sleep signals of full-term neonates. The EEG-sleep data are not stationary.

Neurophysiologists have traditionally identified four distinct electroencephalogram patterns during sleep which we aggregated to just two sleep stages: active and quiet. At first we separated the time series into quasi-stationary increments using change point detection algorithm. For the homogeneous inc

Clustering Trajectories In The Presence Of Informative Patterns Of Monotone Missingness

✦Gabrielle Flynt, Carnegie Mellon University, 5000 Forbes Avenue, Department of Statistics, Baker Hall, Pittsburgh, PA 15213 US, *gfijas@stat.cmu.edu*; Howard Seltman, Carnegie Mellon University; Rebecca Nugent, Carnegie Mellon University; Joel Greenhouse, Carnegie Mellon University

Key Words: Trajectory Analysis, Pattern Mixture Modelss, Missingness, Clustering

Growth mixture models are a method for analyzing longitudinal data that have been recognized for their usefulness in identifying homogeneous subpopulations within the larger heterogeneous population. Missing data is an inevitable obstacle present in longitudinal studies. Missingness is most often dependent on some unobserved variables and may cause biased estimates in statistical procedures that do not account for the informative missing values. Pattern mixture models are an approach to missing data analysis that models the drop-out mechanism. A common assumption used in pattern mixture models is known as the complete case missing variable restriction which uses fully observed trajectories to make parameter estimates for trajectories with missing values. Attempting to cluster growth trajectories without accounting for informative missingness can easily lead to misclassification. The goal of this work is to combine pattern mixture models and trajectory classification while investigating the validity of the complete case missing variable restriction. Results will be shown for several simulated data sets as well as a data set that measures clinical depression in subjects.

Visualize The Changes In Choropleth Maps Through Tc Maps

✦Chunling Zhang, George Mason University, Chicago, IL 60614, *clzhang00@yahoo.com*; Daniel Carr, George Mason University

Key Words: maps, small multiples, change blindness

Looking for and thinking about changes in juxtaposed choropleth maps is complicated by change blindness. Every time our eyes jump in movements called saccades we are blind and our change detectors are reset. All that can be compared is a little area of focal attention in the first map that is sufficiently retained in visual working memory for comparison against the same little area of the second map. Careful comparison requires scrutiny and takes time. Showing changes explicitly as well as the original maps for context provides a general approach that addresses the challenge. The general approach is relevant to sequences of small multiple maps indexed by time or other variables such as sex or age group. However, there are many specific ways to define and represent changes. The methods we developed are implemented in dynamic Java software called TCmaps. This imports boundary files for maps and data for the map regions, and provides a variety of ways to represent data for index series of maps and control the map layouts that show both change and the original maps context.

A Non-Linear Mixed Effects Model Approach To Monotonically Constraining The Parameter Curve In A Functional Non-Linear Model

✦Eduardo Montoya, California State University, Bakersfield, , *emontoya2@csub.edu*; Wendy Meiring, Department of Statistics and Applied Probability, University of California, Santa Barbara

Key Words: functional data, monotonicity, mixed models

Functional regression models that relate a set of functional predictors to a set of scalar responses are becoming more common due to the availability of functional data. A functional non-linear model with a scalar response where the parameter curve is monotone is introduced. Sufficient conditions for casting the functional non-linear model as non-linear mixed effects models are given, providing a monotone parameter curve estimator in a mixed model framework with restricted maximum likelihood for smoothing parameter selection. As an alternative method for estimation, we also present a penalized least squares criterion with generalized cross validation for smoothing parameter selection. Simulation is used to assess the relative efficiency of the monotone parameter curve estimator.

529 Contributed Oral Poster Presentations: Section on Statistical Consulting

Section on Statistical Consulting

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Leveraging Large Transactional Databases Toward Retail Analytics

✦MURLIDHAR JUTTI, PRECIMA, SUITE 600 438 UNIVERSITY AVE, TORONTO, ON M5G 2L1 CANADA, *mjutti@precima.com*

Key Words: SAS, Retail, Analytics, Segmentation, Lifesyle, Lifestage

As the economy grows tighter and competition gets fierce and the customer gets more demanding than ever, increasingly retailers are coming under the microscope to do things differently than they did from the past. Traditional category centric planning has given way to customer centric planning with due emphasis of segmenting the customers based on their lifestyles and lifestage. With the availability of rich databases and sophisticated statistical softwares such as SAS, retailers are now in a position to better understand their customer's needs and wants and can plan their strategies accordingly. Also they are better positioned to face competition from other players and come up with appropriate tactical and strategic moves. In this poster, several case studies will be highlighted to demonstrate how business analytics has evolved into a modern day fact based decision making vehicle that no retailer can stand to ignore.

Statistical Analysis For Electro-Mechanical Response Of Poly Vinyl Alcohol (Pva) And Poly Acrylic Acid (Paa) Hydrogels

✦Ana Tehrani, University of Central Oklahoma, , *stevalle@hotmail.com*

Our objective is to determine the most appropriate statistical method to analyze data collected in biomedical engineering research that involved the shape change characteristics of ionic hydrogel under different electrical stimuli. The hydrogel samples were prepared by dissolving PVA and PAA in deionized water at 4% and mixed together at a 1:1 ratio. Samples were dehydrated then immersed in DI water. Swelling was then recorded in relation to time and electric field which was gradually increased. The rectangular shape change due to the electric field was measured using an optical microscope. The experiment was conducted with the intent of observing stimuli-responsive swelling behavior of the hydrogel which is being investigated for the enhancement of cell function and cell dispersion.

An Application Of Extended Cox Proportional Hazard Approach For Survival Analysis For Recurrent Events

✦Abdul Salam, NGHA, P.O.Box 2477, Al-Hasa, International 31982 Saudi Arabia, *salamab@ngha.med.sa*; Abdul K. Hussein, University of Windsor; Peter Senior, University of Alberta, Clinical Islet Transplant Program,; James Shapiro, University of Alberta, Clinical Islet Transplant Program

Key Words: Type I Diabetes, Islet Transplant, Recurrent event, Extended Cox-proportional model, Log rank test, Kaplan Meier

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 to first event (TTFE), or time to graft failure post last transplant (TTGFPLT). Consequently imprecise estimates of graft survival result from this failure to consider multiple transplants. The objective is to compare extended Cox-based model with classical method for graft survival. Type I diabetes patients who received multiple islet transplants between March 1999 and December 2006 were used to examine the relationship between pre transplant panel reactive antibodies (PRA) on graft survival. Pre transplant PRA of >15% was highly associated with graft failure. Hazard ratio (95% CI) were obtained from Cox-based models for recurrent event AG: 4.3 (1.8 -10.4), PWP-TT: 4.7 (2 -11), PWP-GT: 5(2.3 -12), LWA: 4.3 (1.9 -10) and using classical model TTFE: 3.92 (0.86 -17.9), TTGFPLT: 4.9 (1.7 -13.8). Extended Cox-based models are potentially important in CIT because estimates of hazard ratio are consistent and precise compared with classical methods.

Statistical Analysis Of Questionnaire Data Based On Supplier Information Of A Large Retailer

✦Andy Mauromoustakos, University of Arkansas, 101 AGRX, Agri Stat Lab, Fayetteville, AR 72701, *amauro@uark.edu*; Phil Crandall, University of Arkansas; Ellen J. Van Loo, University of Arkansas; Irina Berdnik, University of Arkansas

Key Words: survey, data, analysis

This current study evaluates the impact of a large retailer requiring of their suppliers to select and become certified under one of the accepted standards by a specific date. The purpose of this study was to document leading indicators, mostly qualitative, that were collected from a representative range of suppliers and distributors as they made the transition. A survey of selected suppliers and their plans showed the majority of the plants to have become compliant and benefited by a significant reduction in the number of audits since they achieved certifications. The statistical analysis and results of the survey will document the tools techniques and treatment of the data collected that will help us to better understand benefits of the various certifications.

An Improved Estimator For The Number Of Landowners Along A Railroad Right Of Way

✦Marcus Nunes, Penn State University, 418 Thomas Building, University Park, PA 16802, *man238@psu.edu*; James Rosenberger, Penn State University; Durland L Shumway, Penn State University

Key Words: sampling, estimation, rails-to-trails, simulation

There are real-life situations where we are interested to know the number of experimental units that share a certain characteristic. In this work, a rails-to-trails organization would like to investigate the likelihood that property owners along an abandoned rail line would cooperate on converting the right of way to a hiking trail. However, finding the number of unique landowners of all segments is time consuming and expensive. The organization proposed a multi-sample procedure, based on the well-known Horvitz-Thompson estimator, to estimate the total number of landowners along the line. The sampling frame is based on a map with probability of landowner selection proportional to the length of each railroad segment traversing a single land parcel. Simulation studies show that this estimator is biased. We investigate another estimator, based on the same samples, which demonstrates better performance. This new estimator, with a bootstrap margin of error, is nearly unbiased, consistent and has lower MSE for estimating the total number of adjacent landowners, which is illustrated with several simulated situations.

Statistical Analysis For Electro-Mechanical Response Of Poly Vinyl Alcohol (Pva) And Poly Acrylic Acid (Paa) Hydrogels

✦Ana Tehrani, University of Central Oklahoma, , *stevalle@hotmail.com*

Key Words: statistical, consulting, undergraduate research, repeated measures

This project is the result of an undergraduate statistical consulting program. An undergraduate statistics student collaborated with an undergraduate student in biomedical engineering on the following research. The objective was to determine the most appropriate statistical method to analyze data that involved the shape change characteristics of ionic hydrogel under different electrical stimuli. Three voltages were compared to a control and the swelling of the hydrogel under each treatment was measured at various time periods. The biomedical engineering student performed the experiment and submitted the resulting data for analysis. A repeated measures analysis of variance was performed resulting in significant differences across time, but no significant differences among the treatments. The results of the biomedical engineering research as well as a description of the collaboration experience will both be presented.

Healthcare Recovery Auditing - Opportunities For Contributions

✦Laura Vazquez, LW Consulting, Inc, 300 Colonial Center Parkway, Suite 100, Roswell, GA 30076, *lvazquez@lw-consult.com*

Key Words: Medical Audit, Standards, Quality, Consulting

In 2006, a 3-year demonstration program to audit payments to fee-for-service medical providers participating in the Medicare program became permanent. Initial geographic limitations were lifted and permanent legislation required expansion to all 50 states by 2010. The purpose of the program is to identify improper payments made to providers for Medicare subscribers. This program allows recovery audit contracts (RACs), working on behalf of the Centers for Medicare & Medicaid Services (CMS), to detect and correct these payments, and implement actions to prevent future improper payments. Subject to the operational and statistical guidelines found in the Program Integrity Manual (PIM), RACs are authorized to use statistical sampling, estimation and extrapolation methods to estimate the improper payment amount. Statisticians should seek emerging opportunities in this area, including development and advocacy for the use of newer, more robust methods in this process; employment and collaboration with auditors to ensure that statistical methods are implemented properly; and utilizing publicly available reports for teaching in the classroom setting.

Where'S The Greek? Applying Diversity Indices To Quantify Achieved Curriculum Breadth

✦Eric Nordmoe, Kalamazoo College, 1200 Academy St, Kalamazoo, MI 49006 USA, *enordmoe@kzoo.edu*

Key Words: diversity index, Shannon's entropy, Gini coefficient, breadth, curriculum

Measures of dispersion for categorical variables have frequently been used to assess diversity in populations. In ecological and economic applications, versions of Shannon's entropy and the Gini coefficient have been adopted as important measures of species diversity and economic inequality respectively. This work applies these and other measures to a college transcript analysis to define and quantify the "breadth" of courses completed by undergraduates over a four-year college record. The value of these measures for comparing differences in achieved breadth across subgroups as well as changes over time is illustrated.

Evaluation Of The Effectiveness Of Two Drugs Combination Therapy In Vitro Cytotoxicity Study

✦Yufeng Li, University of Alabama at Birmingham, 1717 11th Ave. South, Birmingham, AL 35205, *yufengli@uab.edu*; Hui-Chien Kuo, University of Alabama at Birmingham ; Choo Hyung Lee, University of Alabama at Birmingham ; Patsy Oilver, University of Alabama at Birmingham ; Donald Buchesbum, University of Alabama at Birmingham

Key Words: Synergy effect, Drugs combination, Nonlinear regression, Combination index

The cytotoxic of two drugs combination can be evaluated to assess whether the combination cytotoxic effects are additive, less than additive (antagonistic), or greater than additive (synergistic). There are several statistical methods available to evaluate these effects with different statistical properties and limitations. We evaluated the following methods using data generated from a collection of breast cancer cell lines. The methods include 1) recommended by Gennings, modeling using a second order response surface model with linear, quadratic, and interaction terms for each of the cell line. A significant interaction term was classed as either synergistic or antagonistic depending on whether the interaction term was negative with more than additive cytotoxicity or positive with less than additive cytotoxicity. 2) The combination index (CI) for dose combinations were calculated based on the multiple drug-effect equation of Chou-Talalay. 3) The synergism effect was further evaluated with concentration-effective curve with nonlinear regression method by Zhao in 2004.

530 Contributed Oral Poster Presentations: Section on Statistical Education

Section on Statistical Education

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Introducing a Case Study Course as a Capstone Course

✦Harshini Fernando, Purdue University North Central, 1401 S US Hwy 421, Westville, IN 46391, *hfernando@pnc.edu*

Key Words: Case study course, Capstone course, Statistics minor/major

This presentation provides details of incorporating a case study course in an undergraduate statistics minor/major as a capstone course. The case study course as the capstone course requires students to utilize key concepts that they have learned in the other statistics courses in their statistics minor/major. Such course will help the students bridge the connection between statistical methods and its potential applications, making them well prepared for their careers. Also the presentation describes in detail the methodology used in construction of such course, the value and insights about the statistics minor implemented at my institution Purdue University North Central, where we introduced a case study course as a capstone course.

Remediating Misconceptions of Confidence Intervals Using Simulations and Visualizations

✦Chong Ho Yu, Arizona State University, 1475 North Scottsdale Road, Scottsdale, AZ 85257, *chonghoyu@gmail.com*; Samuel DiGangi, Arizona State University; Angel Jannasch-Pennell, Arizona State University

Key Words: Confidence interval, visualization, simulation, Mathematica

Since American Psychological Association released the report on Statistical Inference compiled by the task force led by Leland Wilkinson in 1999, more and more researchers have been reporting confidence intervals (CI) to compensate for the shortcomings of p values yielded from hypothesis testing. However, some researchers (e.g. Leech, Morgan, Wang, & Gliner, 2010; Siegfried, 2010) assert that the same problem of hypothesis testing continues to exist in the usage of CI, and indeed misconceptions of CI are widespread. In order to verify the preceding claims, this study utilizes both objective tests and interviews to identify the common misconceptions of CI using a sample of graduate students in a metropolitan Southwest University. A visualization and simulation program running on Mathematica is used as a remedial tool to test the conjecture that some of the misconceptions of CI could be clarified when the users could literally see the results. A 2X2 design with a between-subject factor (control, treatment) and a within-subject factor (pretest, posttest) is adopted.

Identify Differentially Expressed Genes in Lesional Versus Unlesional Skin Pairs of Psoriasis Patients Using a Meta-Analysis

✦Suyan Tian, The Rockefeller University, New York, NY 10065, *stian@rockefeller.edu*; Mayte Su·rez-FariÒas, The Rockefeller University

Key Words: Meta-analysis, Microarray, Psoriasis, Differentially expressed genes

Psoriasis is a common chronic inflammatory skin disease. Its causes are not fully understood although a genetic component is believed as a must. Expression profiling using microarrays has been widely used to identify genes associated with psoriasis pathology. It is commonly observed that the resulting set of differentially expressed genes (DEGs) differs in different experiments. Therefore, a statistically based meta-analytic approach to combine the results of individual studies is in demand. Here, a random effect meta-analysis on 5 microarray data sets including 193 lesional and non-lesional pairs was conducted. Of 5534 genes used the analysis, 696 (12.5%) were up and 456 (8.2%) were down expressed. Meta-analysis only failed to identify one gene (EPHX3) in the intersection of DEGs upon individual studies, which consisted of 79 up and 22 down expressed genes. However, this gene was filtered out by Integrated Correlation (IGC) method that aims at finding coherently behaved genes across studies. In conclusion, meta-analysis would produce a pool of good candidates of revealing psoriasis pathology and identifying potential treatments.

Multi-Core And Distributed Processing Using R

✦Ian C Crandell, Cal State University, Eastbay, 25800 Carlos Bee Boulevard, Hayward, CA 94542, *icrandell85@gmail.com*; Eric Suess, Cal State University, Eastbay

Key Words: R, multi-core processing, distributed computing, snow, snowfall

Computers have revolutionized applied statistics. As the speeds of processors begin to reach theoretical limits, performance has been improved by the usage of multi-core and distributed processing. The principle behind both multi-core and distributed processing is the same: break up a task into several smaller tasks and execute them on multiple processors in parallel. While this technique can yield great improvements in performance, they can often be hard to implement. This paper is to provide techniques for implementing such forms of processing that can be understood and utilized by a wide audience of statisticians. In order to reach the largest possible audience I will use the free statistical software R. The first part of the submission will be about multi-core processing on a single machine using multi-core and snowfall. The second part will be about linking a network of computers together using snow and then using R across all of them. Using bootstrapping as an example, I will provide code to be run on a single processor, on multiple processors, and across a network.

Developmental Mathematics Into Elementary Statistics

✦Mary R Parker, Austin Community College, Northridge Campus, 11928 Stonehollow Drive, Austin, TX 78758, *mparker@austincc.edu*

Key Words: education, teaching, elementary statistics

The Statway project is a national project to provide a one-year experience to get students to and through a solid Elementary Statistics course, even if their math background is weak. It integrates developmental math topics with activities involving data to help students increase their conceptual skills appropriately so that they can confidently learn the material in a standard Elementary Statistics course. This presentation focuses on what math topics we chose, why, and how we structured the material to have most of the math topics arise in the context of modeling.

Performing One-Group Hypothesis Testing Using Analysis Of Variance

✦J. B. Orris, Butler University, 4600 Sunset Ave., Indianapolis, IN 46208, *orris@butler.edu*

Key Words: hypothesis testing, analysis of variance, ANOVA

In most introductory statisitcs courses and textbooks, hypothesis testing is introduced for one and two group tests and analysis of variance is used as a tool for comparing three or more groups; however, some textbooks and instructors also show that ANOVA also works for comparing two independent groups (One Factor ANOVA) or paired observations (Randomized Blocks). This paper goes a step further and shows how analysis of variance can also be used for one group hypothesis testing (mean vs. hypothesized value). The procedure described is a simple use of randomized blocks ANOVA and it is useful for illustrating the relationships between the different types of hypothesis tests and ANOVA.

Zebras Vs. Hats: Exploiting The Lexical Ambiguity Of The Word Random

✦Neal T. Rogness, Grand Valley State University, Statistics Department, MAK A-1-160, Allendale, MI 49401, *rognessn@gvsu.edu*; Diane G. Fisher, University of Louisiana at Lafayette; Jennifer J Kaplan, Dept. of Statistics and Probability

Key Words: Undergraduate Learning, Language, Word Use

Words that are part of everyday English and used differently in a technical domain possess lexical ambiguity. Because people connect what they hear to what they already know, the use of a common English word in statistics may encourage students to incorporate the statistical usage as a new facet of the word they had learned previously. The use of words with lexical ambiguity, therefore, may encourage students to make incorrect associations between words they know and words that sound similar but have specific meanings in statistics that are different from the common usage definitions. One word integral to the understanding of statistics that has been shown to have lexical ambiguity for undergraduate students is random. Researchers in other fields suggest that in order to help students learn vocabulary instructors should exploit the lexical ambiguity of the words. Our poster will show the results of a study that is part sequence of studies designed to understand the effects of and develop techniques for exploiting lexical ambiguities of random in the statistics classroom and will include class activities and results instructors might obtain from the use of such activities.

Computation Plots For Basic Statistics In R

✦Frank Joseph Matejcik, South Dakota School of Mines, Industrial Engineering, 501 East Saint Joseph St., Rapid City, SD 57701 USA, *frank.matejcik@sdsmt.edu*

Key Words: plots, R, Ancient Tools

Plots developed in XLISP-STAT a few years ago are now available in R. These plots demonstrate the sequence of computation of basic statistics including the mean and standard deviation (from data and from pmf definitions) , the correlation coefficient, the F-test in one way ANOVA, and Chi-squared test for simple contingency tables. These plots may be valuable as alternative teaching presentations. The plots are relatively untested a diagnostics.

Including Student Ability To Assess Learning With Other Assessment Tools

✦Julia Anne Norton, California State University East Bay, 28022 El Portal Drive, Hayward, CA 94542, *julia.norton@csueastbay.edu*; Zachariah Espe Dietz, Hamilton University; John Lovell, California State University East Bay

Key Words: assessment, learning, teaching, introductory statistics

Retrospective look at ten years of assessing introductory statistics courses over quarters, grades, and the introduction of a "how sure are you of this answer" question. Since Fall Quarter 1999, the authorsa have collected data from a required a common final in introductory statistic courses. After ten years we wonder whether there is some relationship between correct response and an individual student assessment of their ability to answer a particular question correctly. Our study considers continuity in exams and the usefulness of asking students to assess their own problem solving ability. For each of twenty questions on a common final in an introductory statistics course, students are asked to rate their personal ability to answer that particular question correctly. Responses are studied on a number of scales. One set of scales is designed to study particular topics in introductory classes. The second set of scales looks at the difficulty of the problems in terms of literacy, skill and reasoning required to answer. In an age requiring "customer satisfaction," we ask whether students are able to correctly utilize basic course skills and assess personal learning.

Credibility Theory In Actuarial Science

✦Ferry Butar Butar, Sam houston State University, 3657 red bud ln, Huntsville, TX 77340, *butar@shsu.edu*

Key Words: credibility, random effect, exponential family, mse, buhlmann

The premium is a weighted average of overall mean and average of the i th group of premium where the coefficient of the mean is called the credible factor. To estimate the credible factor, the model can be developed from standard statistical model such as random effect model. The credibility will be exact if one utilizes the exponential family.

Developing Student'S Intuition And Understanding Of Genetic Variability

✦Mirian Souza, Instituto de Matematica e Estatistica da Universidade de Sao Paulo Brazil, Rua Lucindo Passos Filho, 127 Jd Almanara, Sao Paulo, International 02865040 Brazil, *miria_sou@hotmail.com*

Key Words: Teaching Statistic, Genetic Variability, Family Data

Using a set of developed of teaching activities based on theoretical and practical activities aimed at developing the student intuition and understanding of genetic variability present in different data sets. This result should be applied in a Statistics course with emphasis on analysis of genetic and genomic data. In all activities the students are guided to think about genetic variability through a set of challenging questions discussed in the classroom. The material includes directed activities in which groups of students are motivated to generate using R software, phenotypic variables for family data abiding by family structure and the Hardy-Weinberg equilibrium (HWE). A sampling of the genome of individuals is done through platforms of molecular markers in which one of the most commonly used classes is the type SNPs (Single Nucleotide Polimorphisms) which are approximately 1 million markers. In this context students are instructed to generate the genotypes of individuals following multivariate normal distribution with different structures for the parameters of mean and variance with the aim of investigating the effect of SNP under different scenarios.

Homework + E-Textbook = Integrated Online Learning Of Statistics

✦Brenda Kirsten Gunderson, Department of Statistics, University of Michigan, 439 West Hall, Ann Arbor, MI 48109 USA, *bkg@umich.edu*

Key Words: integrated learning, online homework, e-textbook, customized teaching material

Students are accustomed to accessing information immediately. So we develop ways to enhance the teaching of Statistics and incorporate technological methods into all aspects of the students' learning environment. This poster will share a new online tool that facilitates creation and grading of homework linked to an e-textbook. The idea is to make the e-textbook a supplement to the homework questions. This prototype is being used in an introductory statistics course with over 1500 students. A bank of customized questions has been created and linked directly to e-textbook content. Students work through homework online, using the links as needed. Submission of the paperless homework is automatic and at a common time. Grading is completed online with the ability to provide tailored feedback quickly. We have seen an increase in average grades and an increase of the buy-in of the e-textbook option as students appreciate the integration of textbook with tailored homework questions. Future plans include embedding mini video hints, tagged to specific homework questions. This tool allows students to build connections between the material they encounter to see the bigger picture.

Variable Selection And Inference In Modeling For The Semiparametric Treatment Effect Estimator

✦Shuai Yuan, NCSU, 2819 Broadwell Dr., raleigh, NC 27606 USA, *syuan@ncsu.edu*

Key Words: Semiparametric treatment effect estimation, Semiparametric treatment effect estimation, Covariates adjustment, Shrinkage variable selector, Oracle property, Multiple-split method

Inference on the effects of treatment on the basis of a primary outcome is the objective of randomized clinical trials. Zhang et al. (2008, Biometrics) propose a semiparametric method to estimate the treatment effect using a covariate adjustment to improve efficiency, which involves modeling the regression of outcome on covariates separately by treatment group. We study the use of modern variable selection approaches, including shrinkage-type and false selection rate methods, to identify important covariates to be included in the treatment-specific regression models. We show that the standard error formula for the treatment effect proposed by Zhang et al. underestimates the true sampling variation in finite samples under these conditions. To correct for this underestimation we consider use of the empirical sandwich formula and demostrate its superior performance both through simulations and theoretically. Our simulation results show that shrinkage type methods improve the efficiency of treatment effect estimation in many practical settings.

531 Contributed Oral Poster Presentations: Social Statistics Section

Social Statistics Section

Wednesday, August 3, 10:30 a.m.–12:20 p.m.

Bullying at School and Cyber-Bullying Anywhere: The Current State and Changes Over Time

✦Simone Anke Robers, American Institutes for Research, 1990 K Street NW, Suite 500, Washington, DC 20006, *srobers@air.org*

Key Words: Bullying, School Safety, Ages 12-18, Cyber-bullying, Fear and Avoidance

This effort uses data from the School Crime Supplement(SCS)to analyze bullying at school and cyber-bullying anywhere.Bullying and being bullied at school has increasingly become a topic of interest for researchers from a variety of disciplines,as well as policy makers,law makers and educators.Research provides evidence for short- and long-term psychological difficulties and social relationship problems for the bullied and bullies.Bullying and being bullied can lead to physical violence behavior and gang related activities.The SCS is the supplement to the National Crime Victimization Survey (NCVS)and has collected data on bullying biennially since 1999.Using a variety of questionnaire items related to the topic of interest,the goal of this work is to provide trend analysis over time and analyze most current SCS data available(2009 data available for the author December 2010).This work is aimed to identify the current state,and changes over time,in the occurrence of bullying at school, cyber-bullying anywhere,and fear and/or avoidance behaviors of victims of bullying.The preferred method of analysis is t-tests and regression using statistical analysis software packages SAS and STATA.

Methods for Estimating Selected Life Table Parameters Using the Gompertz Distribution

✦Peter Pflaumer, Technical University of Dortmund, Dortmund, Germany, *peter.pflaumer@uni-dortmund.de*

Key Words: Mortality, Life Table, Life Expectancy, Life Span, Old-Age Dependency Ratio, Net Reproduction Rate

Since the mortality of modern developed population is largely the mortality of old age, the Gompertz model provides a good approximation of life tables in these populations and can be used to estimate and forecast many parameters of the life table and the corresponding stationary population. Average and median remaining life expectancies are calculated and compared using the current U.S. life table. An analytical expression for the median remaining life expectancy is derived. The maximum life span is regarded as a random variable. The median and the modal value of the resulting density function are considered as realizations for the oldest age. A simple method for approximating the old-age dependency ratio is proposed. A sensitivity analysis is carried out. The influence of increasing life expectancy and decreasing net reproduction rate on the old-age dependency ratio is investigated.

Improving Use Of The 3-Parameter Logistic (3-Pl) Model In Item Response Theory Applications

✦Jerry L Gorham, Pearson, Inc., PO Box 72, 82 Rancho de Shama, Box 72, Cerrillos, NM 87010 United States, *jlg8989@yahoo.com*

Key Words: logistic, IRT, likelihood, item response theory, exponential

In this paper I show that the 3-PL item response theory model, unlike the 1-PL or 2-PL models, is not an exponential family model because of the presence of the non-zero lower asymptote or "pseudo-guessing" parameter. It is known that under some conditions the likelihood function may be problematic for latent ability estimation. Traditionally, the issue has been cast in terms of the potential for multiple or non-definite maxima in the likelihood function and methods to ensure that the function will produce a unique estimate. This paper suggests that multiple or non-definite maxima are symptomatic of a more fundamental property of the model - the tendency for misfitting response vectors to distort the likelihood function, especially near the lower end of the scale. This model property can be considered an advantage, in some respects, over the 1-PL or 2-PL models, which remain convex regardless of the degree of misfit. I prove the theoretical conditions under which the likelihood function will be quasi-log concave and provide simulation results to recommend conditions under which the model may be used appropriately.

Writing Errors From Speakers Of African American English, Speakers Of Esl, And Speakers Of Standard American English: Which Annoys Raters More?

✦Lewis VanBrackle, Kennesaw State University, 1000 Chastain Road, Mailbox 1601, Kennesaw, GA 30144, *lvanbrac@kennesaw.edu*; David Johnson, Kennesaw State University

Key Words: logistic model

Raters of state-mandated college-level writing exams are trained to grade holistically when assessing these exams. Such exams are intended to ensure that students possess a minimal writing competency at the university level. A guiding principle in holistic grading is to not focus exclusively on any one aspect of writing but rather to give equal weight to style, vocabulary, mechanics, content, and development. However, the present study, based on 358 ratings of a state writing exam, indicates that raters give more emphasis to surface grammar errors than other areas of writing. In addition, this study details how raters react to surface errors typical of speakers of African American English, ESL, and Standard American English. Using a logistic model to generate odds ratios for comparison of essays with errors typical of these three language groups, results indicate a bias against African American errors and a bias for ESL errors.

New, Testable, Definitions Of Reliability And Measurement Error

Futoshi Yumoto, American Institutes of Research; ✦Rochelle Elaine Tractenberg, georgetown university medical center, Neurology, Suite 202 Building D, 4000 reservoir Rd., NW, washington, DC 20007, *rochelle.tractenberg@gmail.com*

Key Words: classical test theory, reliability, measurement error, psychometrics, cognitive testing

“Reliability” under classical test theory (CTT) assumes that measurement error is: a) identically distributed for all respondents; and b) independent of true score. Neither is testable. A new definition of reliability and measurement error fits the Guttman model to the ability to respond to each item -demonstrated with a cognitive instrument - over time. Incorrect answers at a visit imply incorrect responses at all successive visits, representing “cognitive loss”. Items with response patterns that are consistent with the Guttman model are “reliable”; “measurement error” is an item yielding an inconsistent signal about the individual’s ability to respond. We modeled the first 4 annual evaluations in: a) cognitively normal (NC, N=149); b) diagnosed with possible or probable Alzheimer’s disease (AD) (N=78); and c) cognitively normal with a later diagnosis of AD (converters, N=133). Of 16 test items, reliable measurement was observed for 10 items in NC, eight in converters, and two in AD. We formalized “reliability” over time and functionally defined measurement error. Measurement error was inconsistent over time and cognitive levels, violating the CTT assumptions for reliability.

Enumerative Or Analytic: Statistical Thinking And Scientific Reasoning In Survey Research

✦ZHI GANG WANG, DEPARTMENT OF NATIONAL DEFENCE, 101 Colonel By Dr, Ottawa, ON K1A 0K2 Canada, *zhigang.wang@forces.gc.ca*

Key Words: Survey Research, enumeration, analysis

In his work, Deming urged researchers to distinguish between two types of survey studies - enumeration and analysis. A failure to appreciate the underpinnings of these two types of studies could lead to misuses of statistics and incorrect statistical conclusions when designing and analyzing surveys with complex samples. This paper reviews differences between these two types of studies and their associated inferential statistical paradigms (design-based and model-based inferences). This paper further discusses potential actions to promote sound statistical thinking and scientific reasoning in complex sample survey research.

Meta-Analysis Of Multivariate Outcomes: A Monte Carlo Comparison Of Alternative Strategies

✦Corina M. Owens, University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620-7750 United States, *cmowens@mail.usf.edu*; Jeffrey Kromrey, Educational Measurement and Research, University of South Florida; Julie Gloudemans, University of South Florida

Key Words: meta-analysis, multivariate, random effects, multilevel models

Models for meta-analysis are based on the assumption that effect sizes are independent of each other. However, multiple outcomes and treatments from the same study make the tenability of this assumption doubtful. Several meta-analytic methods have been proposed to handle dependent effect sizes: a multivariate multi-level approach (Kalaian & Raudenbush, 1996); Hedges, Tipton, and Johnson’s (2010) robust variance estimation strategy; and Gleser and Olkin’s (2009) stochastically dependent effect size approach. This research used Monte Carlo methods to compare these approaches to a traditional univariate random effects approach (Hedges & Olkin, 1989). Factors investigated included (a) number of studies included in the meta-analysis, (b) population mean effect sizes, (c) covariance between the effect sizes, (d) within study sample size, and (e) population effect size variance. Accuracy and precision of 95% confidence intervals and bias of the population mean effect size point estimate were investigated. In general, methods produced point estimates with little bias, but provided better coverage when population effect sizes were small and primary study sample sizes were large.

Methods For Designing Fidelity Measure Data Collection For Intervention Studies

✦xin sun, Southern Methodist University, Dallas, TX 75206, *xsun@smu.edu*

Key Words: Effect size, Imperfect fidelity measurement, Measurement error model

Fidelity of implementation in educational intervention study is the extent to which theoretically meaningful components of an intervention are realized in practice. Study shows that level of implementation affects the outcome, so we build a model linking fidelity to outcome and redefine effect size. Asymptotically unbiased estimators of effect size are developed, along with an expression for their standard errors. We present a method for determining an efficient design for collection of fidelity data in intervention studies. The design is optimized based on minimizing variance of estimators of effect size that incorporates fidelity measurement for a given cost.

Evaluation Of Alternative Propensity Score Based Causal Inference Approaches

✦Bing Yu, University of Toronto, 30 Charles St. West, 1407, Toronto, ON M4Y 1R5 Canada, *bingyu@gmail.com*; Guanglei Hong, University of Chicago

Key Words: bias, causal inference, confidence interval, MSE, prognostic score, variance

Extant research on variable selection for propensity score (PS) models selects covariates to be included in a PS model based on their relationships with treatment and/or outcome. In general, including outcome predictors in a PS model improves the precision of treatment effect estimation without increasing bias. This study extends this line of research by incorporating three further aspects. Using simulation data, first we will evaluate the relative effectiveness of combining propensity score with prognostic score adjustment or with strong outcome predictor adjustment. Secondly, we will consider a Normal distributed outcome versus a Bernoulli distributed outcome. Finally, we will test the robustness of alternative approaches under a range of model misspecifications, including omitted covariates, omitted nonlinear terms, or omitted interaction terms in a PS model, a prognostic score model, or an outcome model. Evaluation criteria include bias, precision, MSE, remaining sample size after stratification, and confidence interval coverage percentage.

Estimating The Benefits Of The Home Line Tenant Hotline

✦Craig A Rolling, University of Minnesota, 313 Ford Hall, 224 Church St. SE, Minneapolis, MN 55455, *crolling@stat.umn.edu*

Key Words: estimation, sampling, nonprofit, housing

HOME Line, a nonprofit Minnesota tenant advocacy organization, has operated a free tenant hotline since 1992. All incoming calls are logged in a database, and HOME Line follows up with a sample of those calls to learn how the tenant's situation was ultimately resolved. The follow-up calls can be treated as a stratified sample, where the stratification is by the caller's primary issue or problem. We use this sample to compute estimates and standard errors for the total amount of money recovered and the total number of evictions prevented by all callers to the hotline since 1992. Particular focus will be given to the growing problem of tenants whose landlords are in foreclosure.

Oil Spill Data Analysis

✦Huiming Song, university of California, Riverside, 3382 Kentucky St, Riverside, CA 92507, *songhuiming@gmail.com*; Musen Wen, university of California, Riverside; Shilong Kuang, KBB

Key Words: Oil Spill, Experiment Design

The Deepwater Horizon oil spill is an oil spill in the Gulf of Mexico began from April 20, 2010 and lasted about three month. The spill causes extensive damage to marine and wildlife habitats as well as the Gulf's fishing and tourism industries. In this paper we what to study the relation between the oil spill and birds and fisheries. We will study the oil spill extents with the temperature and salinity. We will show the spatiotemporal pattern of oil presence in the temperature and salinity measurements.

Data Sparseness And Multilevel Models: The Impact Of Small Cluster Size On Point And Interval Estimates In Two-Level Logistic Models

✦Bethany Ann Bell, University of South Carolina, College of Education, WD 133, 802 Main Street, Columbia, SC 29208, *babell@sc.edu*; Jason Schoeneberger, University of South Carolina; Jeffrey Kromrey, Educational Measurement and Research, University of South Florida; John Ferron, University of South Florida

Key Words: multilevel logistic models, sample size, binary predictors, survey research

Methodological research continues to address the important issue of sample size at each level when estimating multilevel models. However, to date, these investigations have primarily focused on multilevel linear models, with normally distributed continuous outcomes. Few studies have examined the impact that sample sizes have on the statistical properties of multilevel models with binary outcomes. Moreover, even less attention has been given to the consequences of estimating multilevel logistic models with sparse level-2 units. To help address this gap in the literature, this Monte Carlo study focuses on the consequences of level-2 sparseness on the estimation of fixed and random effects coefficients in terms of model convergence and both point and interval estimates as a function of the level-1 sample size, number of level-2 units, proportion of singletons (level-2 units with one observation), collinearity, outcome prevalence, intraclass correlation, and model complexity. SAS IML was used to simulate 1000 data sets across 5760 conditions. Results are presented in terms of statistical bias, confidence interval coverage, statistical power, and rates of model non-convergence.

Interpeting Adjusted Means In Analysis Of Covariance

✦Dan Mundfrom, New Mexico State Univ., Las Cruces, NM 88003 USA, *mundfrom@nmsu.edu*; Dennis Clason, New Mexico State University

Key Words: ANCOVA, adjusted means, covariance

In many data analytical applications involving the comparison of means from several populations, analyses are adjusted by including one or more continuous variables in the model. Such an analysis is often called an Analysis of Covariance. In this paper, we examine the effect of such an adjustment on the precision of inferences regarding the adjusted means and the interpretation of those means. We consider the variance of the difference between a pair of adjusted means and show that the variance is minimized when the mean of the covariate is the same in both groups. However, when the covariate means differ across groups, the additional term in the model can dominate the variance of the comparison. We investigate the break-even point for this comparison and under what condition that occurs. Finally, we demonstrate with several data sets, how these results may be manifested in practice.

Sexual Behavior: A Multivariate Analysis Of Forced Threatened Sex Among College Students

✦Leona L Ding, Social and Economic Sciences Research Center, Washington State University, PO Box 644014, Pullman, WA 99164-4014, *leonad@wsu.edu*; Danna Moore, Washington State University

Key Words: sexual behavior, surveys, mutivariate analysis, college students

Sexual violence is a widespread problem among college communities. The purpose of this study is to gather information the pervasiveness and outcomes for students experiencing forced threatened and unthreatening sexual experiences on campus and its outcomes. The source of information for this study was an online survey conducted on a random sample students at the Washington State University. A mixed postal letter invite and web survey protocol obtained over 1,800 students responses. The tentative results indicated that almost a third of female students had experienced unwanted sexual advances and 10% indicated unwanted sexual contact. Most of those experiencing unwanted or threatened sex knew the person involved in the incident. These analyses probe further looking at other personal and environmental characteristics associated with student vulnerability such as living environment and class. Even though the majority (76%) of the students indicated that sexual abuse was a problem between college students today, 85% of them seldom or never discussed sexual assault with their peers and less than 1% filed civil or criminal charges after the incident.

Ordinal Models For Student Evaluation Data

✦Natalie Blades, Brigham Young University, Department of Statistics---230 TMCB, Provo, UT 84602, *blades@stat.byu.edu*

Key Words: ordinal response, bayesian probit model

We propose a context within which student evaluations of teaching may be more meaningfully interpreted. Based on one-million student evaluations, we demonstrate the effects of putative confounders of student ratings and propose a probit model for evaluating instructor effectiveness given demographic information about the students in the course.

532 Section on Health Policy Statistics Speaker with Lunch (fee event)

Section on Health Policy Statistics

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Calibrated Bayes, Models And The Role Of Randomization In Surveys And Experiments

✦Rod Little, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109 USA, *rlittle@umich.edu*

Key Words: statistical inference, likelihood principle, randomization

The calibrated Bayes approach to statistical inference is Bayesian for the inference, but seeks models that yield inferences that are calibrated, in the sense of having good repeated-sampling properties. The Bayesian approach to inference was historically regarded as not supporting randomization for sample selection or treatment allocation, since the randomization distribution is not the basis for inference. I discuss why randomization is important to me as a calibrated Bayesian, and provide some supporting examples

533 Biopharmaceutical Section P.M. Roundtable Discussion (fee event)

Biopharmaceutical Section

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Practical Use of Historical Controls in Clinical Trials for Rare Diseases

✦Jeongsook (Jessica) Kim, US FDA/CBER, 1401 Rockville Pike, Rockville, MD 20852, *jessica.kim@fds.hhs.gov*

Key Words: historical control, rare disease, meta-analysis, clinical trial

In certain clinical trials, one may encounter situations where no active controls are available. In designing such studies, both placebo controls and historical controls have been used as comparators. Understanding historical control rates of some event may present a challenge to the drug development process due to innovations in medical treatments or changes in medical practice over time and the comparability of the specific target patient populations. We will discuss statistical considerations in determining appropriate historical controls for such situations.

Globalization Of Clinical Trials: The Development Of Treatments And Preventative Products For Diseases, With A Focus On Vaccines

✦Tammy Massie, Food and Drug Administration, 1401 Rockville Pike, HFM-215, Rockville, MD 20852 USA, *Tammy.Massie@fda.hhs.gov*

Key Words: Globalization, Clinical Trials, Vaccines

Pharmaceutical development has become a truly global phenomenon as more clinical trials are run in multiple countries & regions. Many issues have emerged-including statistical, clinical, regulatory, logistical, and ethical-reflecting the complex and multi-faceted nature of the problem and the often competing goals and objectives. This session will discuss issues and challenges associated with global vaccine development, but will be applicable to other regulated products including biologics, drugs & medical devices. Clinical studies may be utilized in investigating and examining preventive and prophylactic treatments for infectious diseases that occur globally, including examples from the current global health crisis of diseases such as malaria, e. coli (travelers diarrhea), H1N1, and many other diseases that affect U.S. and non-U.S. residents. Issues that commonly occur in these diseases include differences in standards of care, differences in baseline health status of subject populations, and even ability to recruit and retain subjects. Ensuring these issues are minimized are critical to pre-clinical to clinical studies.

Regression With Latent Variables (Mimic Models): A Better Way To Analyze Composite Scores From Instruments In Clinical Trials And Medical Research

✦Chengwu Yang, Pennsylvania State University College of Medicine, A210, ASB 3400H, 600 Centerview Drive, Hershey, PA 17033, *yangc@psu.edu*; Barbara C. Tilley, University of Texas Health Science Center at Houston ; Anbesaw Selassie, Medical University of South Carolina; Ruth Greene, Johnson C. Smith University

Key Words: Latent Variables, Subjective Outcomes, Patient-Reported Outcomes (PRO), Multiple-indicator Multiple-causes (MIMIC) Models, Differential Item Functioning (DIF), Structural Equation Modeling (SEM)

Subjective outcomes (e.g., Quality of Life) are prolifically used in clinical trials and medical research, and they are measured by composite scores from instruments (e.g, SF-36). The instruments' validity is a big concern, and an important validity issues is if the instrument's factor structure is sustained. The conventional multivariate regression that directly regresses the composite scores on important covariates cannot investigate the factor structure, and therefore the results can be wrong. The multiple-indicator multiple-causes (MIMIC) model that regresses the latent domain scores on the same important covariates has these advantages: 1) it can assess factor structure of an instrument; 2).it can investigate if any covariate effect on the composite scores is contaminated by measurement bias, i.e., differential item functioning (DIF). Three situations exit when applying MIMIC models to composite scores: 1) the instrument's factor structure sustained and there is no DIF; 2) the instrument's factor structure sustained but there is DIF; 3) the instrument's factor structure didn't sustain. Appropriate analyzes strategies corresponding to these three conditions will be discussed.

Identifying Treatment-Effect Heterogeneity In Clinical Trials Using Subpopulation Treatment Effect Pattern Plots

✦Ann A Lazar, University of California, San Francisco, 3333 California Street Ste 495, San Francisco, CA 94119 USA, *ann.lazar@ucsf.edu*; Wai-ki Yip, Harvard School of Public Health; Bernard F. Cole, University of Vermont; Marco Bonetti, Bocconi University; Richard D Gelber, Harvard School of Public Health and Dana-Farber Cancer Institute

Key Words: Competing Risks, Biomarkers, Kaplan-Meier, Breast Cancer, Hazard Ratio, Interaction

This session will focus on how to identify whether a subgroup of patients responds differently to a treatment than the overall population. In particular, we will discuss Subpopulation Treatment Effect Pattern Plots (STEPP), an intuitive graphical method designed to evaluate treatment-effect heterogeneity. The results from STEPP enhance the interpretation of clinical trial results needed for making individual patient decisions. Examples from a Phase III clinical trial will be included along with the associated code from the R STEPP software package.

Adaptive Designs In A Confirmatory Setting

✦Jeff Maca, Novartis Pharmaceuticals Corporation, One Health Plaza, East Hanover, NJ 07936-1080, *jeff.maca@novartis.com*

Key Words: Adaptive designs, Clinical Trials, Confirmatory Studies, Sample Size reestimation

Over the last few years, there has been much research, and now much implementation of adaptive designs in clinical trials. With the release of the FDA guidance document, many people are implementing, or planning on implementing such clinical trial designs. Now would be a good opportunity to discuss what has been learned and what issues which must be addressed to properly conduct these types of study designs.

Current Issues In The Design And Analysis Of Non-Inferiority Trials

✦Gang Li, J&J Pharmaceutical Research and Development, L.L.C., 920 US Highway 202 S, Raritan, NJ 08869, *GLi@its.jnj.com*

Key Words: Inferiority index, non-inferiority margin, bioequivalence, biosimilarity

Despite the release in 2010 of an FDA draft guidance on non-inferiority trials, controversies still exist. Perhaps the most difficult issue in designing non-inferiority trials has been specification of the non-inferiority margin, particularly when there's insufficient historical placebo-controlled information. The subjectivity of the currently used approaches for margin specification is also a source of the controversies. This session is to provide a platform for further dialogue on non-inferiority trials, including a new approach for margin specification.

534 Business and Economic Statistics Section P.M. Roundtable Discussion (fee event)

Business and Economic Statistics Section

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Electronic Commerce, Data, and Statistical Challenges

✦Wolfgang Jank, University of Maryland, 4322 Van Munching Hall, College Park, MD 20742-1815 USA, *wjank@rhsmith.umd.edu*

Key Words: internet, dynamic models, data mining

Internet transactions create huge floods of complex and challenging data. Every time we purchase an item on Amazon, post a message on Facebook or search on Google, someone records our every click and move of the mouse. This wealth of information creates challenges but also opportunities for statistical innovation. Information changes dynamically and requires real-time predictions. Users are geographically vastly dispersed and communicate in a variety of different languages - yet statistical algorithms have to understand what a user wants, when she wants it and how urgently. Statistical challenges include extracting spatio-temporal differences based on text, voice or relational information. In this round table, we intend to discuss some of these challenges as well as opportunities for Statistics to get involved.

535 Section for Statistical Programmers and Analysts P.M. Roundtable Discussion (fee event)

Section for Statistical Programmers and Analysts

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Producing Publication-Quality Graphs

◆Nancy J. Petersen, Department of Veterans Affairs, Houston HSR&D Center of Excellence, 2002 Holcombe Blvd. (152), Houston, TX , *petersen.nancyj@va.gov*

Key Words: statistical software, graphs, plots

Publication-quality graphs are often needed when results of research studies are submitted to journals. Investigators may rely on statistical programmers and analysts to produce these graphs. This roundtable will discuss which software packages would be best for producing graphs such as trend lines, survival curves, ROC curves, and other common plots.

Unleashing The Power Of The Scientific Programmer In The Pharmaceutical Industry

◆Amy E Gillespie, Merck & Co., Inc, P.O. Box 1000, Upper Gwynedd, PA 19454-2505, *amy_gillespie@merck.com*; Lori Mixson, Merck

Key Words: statistical programming, early phase, epidemiology, economics, SAS, R

Statistical programmers in the pharmaceutical industry primarily focus their support to phase II and III clinical development research, utilizing SAS to develop analysis and submission ready datasets, tables, and figures. Our roundtable discussion will challenge this traditional viewpoint as we debate strategic ways of utilizing the skills of statistical programming professionals in other areas of drug development. Topics for discussion will include statistical programming in support of early phase development, epidemiology, health economics, challenges of working with novel data types and requirements to preprocess, filter, and manipulate data prior to analyses, development of standards for areas that require flexibility and exploratory environments, and software usage including SAS, R, Matlab, and Splus.

536 Section on Bayesian Statistical Science P.M. Roundtable Discussion (fee event)

Section on Bayesian Statistical Science

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Rich-Data Problems in Econometrics

◆Hedibert Freitas Lopes, The University of Chicago, 5020 South Woodlawn Avenue, Chicago, IL 60637, *hlopes@chicagobooth.edu*

Key Words: Variable selection, Factor regression for forecasting, Large Vector Autoregressions

Rich-data problems are becoming commonplace in many areas of sciences such as molecular biology and evolution, environment and climate modeling and forecasting, to name a few. The same trend is observed in econometrics, particularly in macroeconomics and financial econometrics. Such rich-data environments have posed challenging and computationally expensive obstacle to direct implementation of standard statistical tools through all steps of modeling, i.e. data processing, statistical inference and summary and forecasting. In this roundtable we will talk about such difficulties, learn how applied econometricians are tackling them.

537 Section on Government Statistics P.M. Roundtable Discussion (fee event)

Section on Government Statistics

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Career Development For Statisticians: Making Creative Use Of Future Opportunities And Challenges

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

Key Words: Balance of statistical quality, cost and risk, Communication, Perceived utility functions, Professional norms and standards, Societal and scientific context, Stakeholders

Over the next decade, all statisticians working with the federal statistical system will encounter an extraordinary range of opportunities and challenges arising from the increasing need for sophisticated statistical information. This roundtable discusses career development for statisticians within the context of these opportunities and challenges, with emphasis on the development and use of skills to do the following. a. Understand the societal and scientific context in which we carry out statistical work. Of special importance are the professional norms and perceived utility functions of primary stakeholders. b. Exercise well-informed judgment on where and how sophisticated statistical methodology can add substantial value to areas in which stakeholders need high-quality information. c. Develop a rigorous understanding of the balance among benefits, costs and risks encountered in development, implementation and operation of the methodology identified in (b). d. Articulate all of the ideas above in ways that are clear, accessible and tuned to the norms and perceived utility functions of the statistical profession, primary stakeholders, and the public

538 Section on Health Policy Statistics P.M. Roundtable Discussion (fee event)

Section on Health Policy Statistics

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Improving Communications Between Fda, Sponsor And Advisory Panel

✦Gary Lynn Kamer, FDA/CDRH/OSB/DBS, 10903 New Hampshire Avenue, Bldg 66/ Room 2224, Silver Spring, MD 20993, *gary.kamer@fda.hhs.gov*

Key Words: Sponsor, Panel, Pre-IDE, Pre-PMA, Communications

This discussion will identify the information the manufacturing sponsor of a medical device needs to provide to FDA to improve the efficiency of the required FDA review of the safety and effectiveness of the device. Some of this information may be useful prior to the sponsor's submission of an original Investigational Device Exemption (IDE). Other information may be necessary only after the completion of the testing and studies required under the IDE, but prior to the submission of a Pre-Market Application (PMA). The PMA describes the implementation of the tests and studies performed, as well as their results. More importantly, the PMA provides the sponsor's justification for the marketing of the medical device, based on it's safety and effectiveness as demonstrated by the analyses of the data. Also, the need for effective communications between the sponsor, FDA and FDA Advisory Panel members will be examined. Examples will be provided.

Relative Efficiency For Various Designs Of Biomarker-Directed Clinical Trials

Xiaofei Wang, Duke University Medical Center; ✦Stephen George, Duke University Medical Center, Durham, NC 27710, *georg001@mc.duke.edu*

Key Words: Biomarker, Biomarker-directed study, Clinical trials, Personalized medicine, Relative efficiency, Target agents

Biomarkers play an important role in disease diagnosis, prognostic stratification and prediction of outcome, and ultimately in personalized medicine. New drugs are increasingly being developed to target patient subgroups with unique molecular signatures. The efficacy and safety of targeted agents and the predictive accuracy of the associated biomarker need to be evaluated in clinical trials. This round table will focus on the statistical and cost issues in the design and analysis of biomarker-directed clinical trials for evaluating biomarker-directed treatment. Common designs in biomarker-directed clinical trials, biomarker-targeted designs, biomarker-stratified designs and biomarker-strategy designs, will be reviewed. The relative efficiency in conducting clinical trials using these designs will be discussed in term of the number of required patients, study time and cost.

539 Section on Quality and Productivity P.M. Roundtable Discussion (fee event)

Section on Quality and Productivity

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Measurement System Assessment

✦Stefan Steiner, University of Waterloo, Waterloo, ON N2L 3G1 Canada, *shsteine@uwaterloo.ca*

Key Words: Gauge R&R, leveraging, intraclass correlation, measurement reliability

In this roundtable, we discuss how to add some excitement to the necessary task of assessing a measurement system. Recent substantive improvements include the use of leveraging (i.e. repeated measuring of extremes), baseline data and non-standard designs. What else is possible? Bring your ideas.

540 Section on Statistical Computing P.M. Roundtable Discussion (fee event)

Section on Statistical Computing

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Speeding Up The R Language Learning Curve For Students And New Users

D. Keith Williams, University of Arkansas for Medical Sciences; ✦Zoran Bursac, University of Arkansas for Medical Sciences, 4301 West Markham, Slot 781, Little Rock, AR 72205, *zbursac@uams.edu*

Key Words: R, New users, Teaching

The R language is becoming increasingly popular within and outside of academic settings. Due to its advantages, for instance being a freeware and growing international users group, students are starting to use R more often as an analytical software of choice, and carrying it forward into a work place. One barrier to its adoption by new users is arguably a steep learning curve. The purpose of this round table is to discuss strategies, solutions, materials and add-ons that can give a learning momentum to new users. Applications of this round table are suitable but not limited to teaching R in introductory (bio)statistics courses to new users as well as SAS/SPSS practitioners, for basic analysis, graphics and beyond.

Smoothing Splines With Product Kernels In Modeling Of Computer Experiments

✦Zhanyang Zhang, Department of Statistics, University of Michigan, 200 S University Ave, University of Michigan, Ann Arbor, MI 48104 United States of America, *zyzhang@umich.edu*

Key Words: Computer experiments, Gaussian Process, Smoothing Splines, Bayesian method, Kernel functions, Krigging

In many scientific studies, computer codes are used to simulate complicated physical processes. These computer codes are usually referred as computer experiments. One typical difficult of running computer experiments is the computational cost, in which one run can take days to finish. Thus there is a crucial need for efficient statistical surrogate models to predict the computational model using limited number of runs. One widely used statistical surrogate is the Bayesian Gaussian Process model, which utilizes the key assumption that the output of a computer experiment is a realization of Gaussian Process. In this work, we tried to relax the Gaussian Process assumption by putting the prediction problem into a regression framework. The proposed method here is called Smoothing Splines with product kernels. Smoothing spline models are flexible nonparametric regression models which are particularly suitable for modeling of computer experiments. In this work, we also compared the commonly used Bayesian Gaussian Process models with the proposed product smoothing spline models and showed that the proposed model is superior in terms of prediction error and model selection.

541 Section on Statistical Consulting P.M. Roundtable Discussion

Section on Statistical Consulting

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Remote Statistical Consulting

✦Charles Kincaid, COMSYS, 5220 Lovers Lane, Portage, MI 49002 USA, *ckincaid@comsys.com*

Key Words: Virtual Consulting, Remote collaboration, Global

Our discipline is all encompassing in many ways, not the least of them is geography. As our world becomes more global and more virtual we more often consult with clients geographically removed from us. Sometimes with people that we've never met face-to-face. We will discuss the tools, techniques, travel, protocols and best practices for how to be a successful remote statistical consultant. Whether you are an independent consultant working from home, collaborating with a colleague in another country, or analyzing manufacturing data from another hemisphere, join us for a challenging and informative discussion.

542 Section on Statistical Education P.M. Roundtable Discussion (fee event)

Section on Statistical Education

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Uses, Abuses, Underuse, and Resistance to Excel in Teaching Statistics

✦John Barroso, Duquesne University, 600 Forbest Ave, Rockwell Building 8th floor Mailbox, Pittsburgh, PA 15282, *barrosoj@duq.edu*

Key Words: Excel, Business Statistics, Statistical Education, Add-in, Graphing, Excel functions

This Round Table is focused at discussing the use of Microsoft Excel in Statistical Teaching in Business, Education, and Humanities in general. Subtopics include a) Excel Practicality versus Graphing weaknesses, b) Underuse of Excel's functions and features in classroom, c) Distribution Tableless books? d) Usefulness and complications with Add-ins, including lab-licenses e)Is "Excel in Statistics" only for Business Statistics? f) "Hardcore Scientists" versus Excel Aficionados: clashes, outcomes, future trends.

Online vs. Brick and Mortar

✦Peter Bruce, statistics.com, 612 N. Jackson St., Arlington, VA 22201 USA, *peter.bruce@statistics.com*

Key Words: online education, introductory statistics

Teaching Statistics Online: Online education is growing faster than brick and mortar classes. What is the future of higher education? How will statistics fare? Does teaching intro stats require face-to-face instruction? Will a "best-of-breed" online version of Statistics 101 develop and come to dominate the field? If so, what will contribute to such an outcome? If not, what will prevent it? These are some of the questions that will be discussed at this roundtable.

Online Computation For Teaching Statistics

✦J.J. Allaire, RStudio, Inc., 581 Boylston Street, Suite 701, Boston, MA 02116, *jj@rstudio.org*

Key Words: R, Online, Web, Tools, RStudio

One of the challenges in using technology when teaching statistics is getting all your students on the same page: licensing and installing the software, coping with different interfaces for different operating systems, distributing data and software files. †A new approach is to use web interfaces so that students need only have a browser. †This roundtable will present one on-line system based on R. †In addition to removing issues about installation and system compatibility, on-line systems provide attractive opportunities for group and collaborative work. †We'll discuss how on-line computation and collaborative possibilities can be used effectively and explore some innovative ways to provide better interfaces between students and statistical computation.

543 Section on Statistics in Defense and National Security P.M. Roundtable Discussion (fee event)

Section on Statistics in Defense and National Security

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Rigor In Defense Testing And Evaluation: Balancing Best Practices And Implementability

✦Catherine Warner, Director of Operational Test and Evaluation, 1700 Defense Pentagon, Washington, DC 20301-1700 USA, *Catherine.Warner@osd.mil*

Key Words: Defense Test and Evaluation, Statistical Rigor

The Department of Defense Test and Evaluation (T&E) community plans, coordinates, and evaluates the testing of complex military systems. This testing is ripe for the application of statistical techniques and analyses. However, as seen in the recent push by the Director of Operational Test and Evaluation to include Design of Experiments (DOE) in T&E, there are educational and expertise limitations in the T&E workforce that limit the application of advanced statistical methods. In this roundtable we will discuss the potential benefits of increasing the use of statistical methods in T&E. Questions to ponder include: What are the potential gains of increased statistical rigor? How much statistical rigor is enough? Does the principle of diminishing returns apply to the application of increasingly complex statistical methods to T&E? What are the educational needs of the T&E work force to increase the use of statistics in T&E?

544 Section on Survey Research Methods P.M. Roundtable Discussion (fee event)

Section on Survey Research Methods

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

The General Social Survey (GSS) and the Methodology for Studying Societal Change

✦Tom W Smith, NORC/University of Chicago, 1155 East 60th Street, Chicago, IL 60637, *smitht@norc.uchicago.edu*

Key Words: Reliability, Validity, Measurent artifacts, Survey methodology, Societal trends, Social indicators

The GSS has been tracking societal trends since 1972. It has measured changes in over 1,000 attitudinal, behavioral, and demographic variables. In carrying out this widespread social indicators monitoring, the GSS has developed both theoretical and applied approaches for maximizing the reliability and validity of trends and minimizing the occurrence of measurement artifacts. We will draw on specific GSS experience to illustrate the do's and don't of studying societal change.

545 Section on Teaching of Statistics in the Health Sciences P.M. Roundtable Discussion (fee event)

Section on Teaching of Statistics in the Health Sciences

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Statistics In Nursing Research

✦Matthew J Hayat, Johns Hopkins University, 525 N. Wolfe St, Office 532, Baltimore, MD 21205 USA, *mhayat2@son.jhmi.edu*

Key Words: nursing research, statistics education, statistical consulting, evidence based practice, psychometrics

Statistics education is an essential component of nursing education. This is a result of the focus in the nursing field on the use of evidence-based nursing practice. Each undergraduate and graduate nursing program requires successful completion of one or more statistics courses. The nursing researcher often makes use of particular statistical methods, including a frequent focus on the use of psychometrics. In this roundtable discussion we will discuss the statistician's role in educating and consulting with nurse researchers.

546 Social Statistics Section P.M. Roundtable Discussion (fee event)

Social Statistics Section

Wednesday, August 3, 12:30 p.m.–1:50 p.m.

Who Is Poor? A New Look With The Supplemental Poverty Measure

✦Kathleen S Short, US Census Bureau, SEHSD, Rm 7H171, Washington DC, DC 20233, *kathleen.s.short@census.gov*

Key Words: poverty, poverty measures, Supplemental Poverty Measure

For many years, there have been critiques of the official poverty measure for the U.S. On March 2, 2010, the Commerce Department announced plans for the development of a Supplemental Poverty Measure (SPM) "that will use the best new data and methodologies to obtain an improved understanding of the economic well-being of American families and of how federal policies affect those living in poverty." This group will discuss many of the elements recommended by the Interagency Technical Working Group to include in a preliminary Supplemental Poverty Measure. Measurement issues, updating methods, and plans for future research on poverty estimates at the Census Bureau will be described and discussed.

547 Recent Avances in Change-point Problems in Econometrics and Statistics ●

Business and Economic Statistics Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Segmenting Mean-Nonstationary Time Series via Trending Regressions

✦Alexander Aue, University of California, Davis, Department of Statistics, One Shields Avenue, Davis, CA 95616 United States, *alexaue@wald.ucdavis.edu*; Lajos Horvath, University of Utah; Marie Huskova, Charles University Prague

Key Words: Change-point analysis, Circular bootstrap, Extreme value theory

We provide a segmentation procedure for mean-nonstationary time series. The segmentation is obtained by casting the problem into the framework of detecting structural breaks in trending and possibly dynamic regression models in which the regressors are generated by suitably smooth (random) functions. As test statistics we propose to use the maximally selected likelihood ratio statistics and a related statistics based on partial sums of weighted residuals. The main theoretical contribution of the paper establishes the extreme value distribution of these statistics and their consistency. To circumvent the slow convergence to the extreme value limit, we propose to employ a version of the circular bootstrap. This procedure is completely data-driven and does not require knowledge of the time series structure. In an empirical part, we show in a simulation study and an application to air carrier traffic data that the finite sample performance is very satisfactory.

Estimating Structural Changes in Regression Quantiles

✦Zhongjun Qu, Boston University, , *qu@bu.edu*; Tatsushi Oka, National University of Singapore

Key Words: Structural change, quantile regression, conditional distribution, Policy evaluation

This paper considers the estimation of multiple structural changes occurring at unknown dates in one or multiple conditional quantile functions. The analysis covers time series models as well as models with repeated cross sections. We estimate the break dates and other parameters jointly by minimizing the check function over all permissible break dates. The limiting distribution of the estimator is derived and the coverage property of the resulting confidence interval is assessed via simulations. A procedure to determine the number of breaks is also discussed. Empirical applications to the quarterly US real GDP growth rate and the underage drunk driving data suggest that the method can deliver more informative results than the analysis of the conditional mean function alone.

A Self-Normalized Approach to Testing for Change Points in Time Series

✦Xiaofeng Shao, University of Illinois at Urbana-Champaign, Department of Statistics, 725 S. Wright Street, Champaign, IL 61820 USA, *xshao@uiuc.edu*

Key Words: change point, cusum, invariance principle, self-normalization

In this talk, we introduce a new class of change point test statistics in the time series setting. To test for a mean shift, the traditional Kolmogorov-Smirnov CUSUM-based test statistic involves a consistent long run variance estimator, which is needed to make the limiting null distribution free of nuisance parameters. The commonly used long run variance estimator requires to choose a bandwidth parameter and its selection is a difficult task in practice. The bandwidth that is a fixed function of the sample size is not adaptive to the magnitude of the dependence in the series, whereas the data-dependent bandwidth could lead to nonmonotonic power. To circumvent the difficulty, we propose a self-normalization (SN) based Kolmogorov-Smirnov test, where the formation of the self-normalizer takes the change point alternative into account. The resulting test statistic is asymptotically distr

548 Learning Network Structures

IMS

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

A New Approach To Cholesky-Based Covariance Regularization In High Dimensions

✦Adam Rothman, University of Minnesota, School of Statistics, 224 Church Street, S.E., Minneapolis, MN 55455, *rothman@stat.umn.edu*; Liza Levina, University of Michigan; Ji Zhu, University of Michigan

We propose a new regression interpretation of the Cholesky factor of the covariance matrix, as opposed to the well-known regression interpretation of the Cholesky factor of the inverse covariance, which leads to a new class of regularized covariance estimators suitable for high-dimensional problems. Regularizing the Cholesky factor of the covariance via this regression interpretation always results in a positive definite estimator. In particular, one can obtain a positive definite banded estimator of the covariance matrix at the same computational cost as the popular banded estimator of Bickel & Levina (2008b), which is not guaranteed to be positive definite. We also establish theoretical connections between banding Cholesky factors of the covariance matrix and its inverse and constrained maximum likelihood estimation under the banding constraint, and compare the numerical performance of several methods in simulations and on a sonar data example.

Penalized Principal Component Regression on Graphs for Analysis of Subnetworks

✦George Michailidis, University of Michigan, 453 West Hall, 1085 South University, Ann Arbor, MI , *gmichail@umich.edu*

Key Words: Penalized regression, graphs, Laplacian

Network models are widely used to capture interactions among component of complex systems, such as social and biological. To understand their behavior, it is often necessary to analyze functionally related components of the system, corresponding to subsystems. Therefore, the analysis of subnetworks may provide additional insight into the behavior of the system, not evident from individual components. We propose a novel approach for incorporating available network information into the analysis of arbitrary subnetworks. The proposed method offers an efficient dimension reduction strategy using Laplacian eigenmaps with Neumann boundary conditions, and provides a flexible inference framework for analysis of subnetworks, based on a group-penalized principal component regression model on graphs. Asymptotic properties of the proposed inference method, as well as the choice of the tuning parameter for control of the false positive rate are discussed in high dimensional settings. The performance of the proposed methodology is illustrated using simulated and real data examples from biology.

Estimation Of Multiple High-Dimensional Directed Acyclic Graphs

✦Yiping Yuan, University of Minnesota, Twin Cities, Department of Statistics, 224 Church St SE, Rm 313, Minneapolis, 55455, *yuan0076@umn.edu*

Directed acyclic graphs(DAGs) are a popular tool to represent causal relationship among random variables. Besides estimation of a DAG, detection of changes in a graph is also of our interest when we have more than one data set from the same random variables. These changes usually reveal important causality facts in the regarding Øeld. In addition, when the ordering of nodes is known, the problem of estimating DAGs reduces to the problem of estimating the structure of the network. In this article, we consider two adjacent DAGs from the same Gaussian random variables with a natural ordering. Sparsity is assumed in the context of high-dimensionality. We propose a method to estimate the adjacency matrices of two DAGs and in theory multiple DAGs with an approximate L0 constraint on the elements of adjacency matrices of DAGs to encourage sparsity and with another approximate L0 constraint on the element-wise diÆerence of adjacency matrices to detect changes in the graph. Computationally, we introduce an e±cient algorithm that is based on the augmented Lagrange multipliers, diÆerence convex method and a fast algorithm for solving the original LASSO problem. Compared to the Lagrange version of the problem, this constrained setting gives better theoretical properties as well as easier tuning for the tuning parameters. We also demonstrate the method for simulated data and real data.

549 Computational and inferential Issues in Spatio-Temporal Modeling ■●

Section on Statistics and the Environment, ENAR, Section on Physical and Engineering Sciences, Section on Statistical Computing, Section on Statistics in Epidemiology

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

A Computational Framework for the Statistical Analysis of Climate Model Ensembles

✦Stephan Sain, NCAR, P.O. Box 3000, Boulder, CO , *ssain@ucar.edu*

Key Words: Climate model ensemble, Climate change, Reduced-rank Kriging, Markov random field

The statistical analysis of climate model ensembles presents a number of challenges, including the size and complexity of the model output in the ensemble. In this talk, I will discuss these challenges within the context of a large-scale, multi-model regional climate experiment. Further, a statistical framework for evaluating models and producing probabilistic projections of climate change will be presented. This approach can be considered a type of reduced-rank Kriging and incorporates a multi-resolution basis expansion with a regularization based on a Markov random field prior distribution on the coefficients. Case studies will include temperature change over North American and detecting the ENSO response in precipitation fields.

Statistical Analysis of Massive Spatio-Temporal Data Sets

✦Michael Stein, University of Chicago, , *stein@galton.uchicago.edu*

This talk will consider various issues relating to the statistical analysis of massive spatio-temporal datasets such as modeling nonstationarity and approximate likelihood computation. Applications may be drawn from the atmospheric sciences, climatology or nuclear engineering and include computer model output or observational data.

Evaluation of Bayesian Spatial-Temporal Latent Models in Small-Area Health Data

✦Jungsoon Choi, Medical University of South Carolina, 135 Cannon Street, Suite 303, Charleston, SC 29425-8350, *choju@musc.edu*; Andrew B Lawson, Medical University of South Carolina

Key Words: Spatial cluster, diagnostic, Spatial temporal mixture model, latent model, small area health

Health outcomes are linked to air pollution, demographic, or socioeconomic factors which vary across space and time. Thus, it is often found that relative risks in spatial health data have locally different patterns. In such cases, latent modeling is useful in the disaggregation of risk profiles. In particular, spatial-temporal mixture models can help to isolate spatial clusters each of which has a homogeneous temporal pattern in relative risks. Mixture models are assumed as they have various weight structures and considered in two situations: the number of underlying components is known or unknown. In this paper, we compare spatial-temporal mixture models with different weight structures in both situations. For comparison, we propose a set of spatial cluster detection diagnostics which are based on the posterior distribution of weights. We also develop new accuracy measures to assess the recovery of true relative risk. Based on the simulation study, we examine the performance of various spatial-temporal mixture models in terms of proposed methods and goodness-of-fit measures. We examine two real data sets: low birth weight data and chronic obstructive pulmonary disease data.

Disease Mapping by Parsimonious Clustering Technique

✦Taps Maiti, Michigan State University, Department of Statistics and Probability, A413 Wells Hall, East Lansing, MI 48824 USA, *maiti@stt.msu.edu*

Key Words: Clustering, Disease Mapping, Relative Risk

Use of GIS data is common in disease mapping. Due to lack of neighborhood definition and other reasons, clustering could be an alternative to Gaussian Markov random field based methods. The standard clustering technique does not produce the area or sub-population specific disease risks. In this talk we present a new technique that can detect the high (low) risk clusters as well as the area specific disease risks.

550 Wald Lecture: Random walks from statistical physics (II) ●

IMS, International Chinese Statistical Association, SSC

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Two Dimensions and Conformal Invariance

✦Gregory F Lawler, University of Chicago, , *lawler@math.uchicago.edu*

It was predicted by physicists that scaling limits of systems “at criticality” in two dimensions exhibit some kind of conformal invariance. There has been a great surge of activity in the last twelve years making much of this precise. The main new tool is the Schramm-Loewner evolution (SLE) but there are other interesting objects as well. This talk will introduce SLE and discuss what can be proved using this.

551 Analysis of Object Data at SAMSI ■●

Statistical and Applied Mathematical Sciences Institute, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Object-Oriented Data Analysis: Trees as Data

✦Steve Marron, Department of Statistics, University of North Carolina, 309 Smith, Chapel Hill, NC 27516, *marron@email.unc.edu*

The relatively new field of OODA has made a lot of progress during the recent SAMSI program. A particularly challenging aspect of this viewpoint, is populations of tree-structured objects. Deep challenges arose, which involve a marriage of ideas from statistics, geometry, and numerical analysis, because the space of trees is strongly non-Euclidean in nature. These challenges, together with some approaches to addressing them, are illustrated using a real data example, where each data point is the tree of blood vessels in one person's brain.

Manifold Data Analysis

✦Ian L Dryden, University of South Carolina, Department of Statistics, LeConte College, Columbia, SC 29208, *dryden@mailbox.sc.edu*

Key Words: Shape, Manifold, Geodesic, Principal components analysis, Sphere, Metric

Data that lie in a curved manifold are increasingly encountered in many application areas, for example in medical image analysis and computer vision. Classical manifold examples include spheres and landmark shape spaces, but newer manifold applications include spaces of curves, surfaces, diffusion tensors, diffeomorphisms and tree structures. Even defining the notion of a mean in a manifold is often not straighforward, and so care must be taken with the development of statistical methodology. Analogies of principal components analysis, regression models and smoothing splines in manifolds are some of the topics of interest. Particular distinctions are made between mildly non-Euclidean data, where the differential geometric structure is available, and strongly non-Euclidean data, where singularities and other complications abound.

552 Are Hurricane Risks Increasing? ■●

Section on Risk Analysis, International Chinese Statistical Association, Section on Statistics in Defense and National Security, Scientific and Public Affairs Advisory Committee, Committee of Representatives to AAAS

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

On the Increasing Intensity of the Strongest Atlantic Hurricanes

✦James Elsner, Florida State University, Tallahassee, FL 32306, *jelsner@fsu.edu*; Thomas Jagger, Florida State University

Key Words: hurricane, climate change, quantile regression, North Atlantic, trend, Gulf of Mexico

The past three decades have seen a significant upward trend in the intensity of the strongest hurricanes worldwide. Questions remain about this trend especially its relevance to coastal communities in the United States and elsewhere. I focus on observed changes in the intensity of the strongest hurricanes over the North Atlantic basin and on the spatial pattern of these changes. Results show that the upward trend is significantly related to rising sea-surface temperature (SST) after accounting for El Nino. The trend peaks at 16 m/s per C at the 75th percentile with a 90% confidence interval of between 7 and 20 m/s per C. The consequences of increasing intensity of the strongest hurricanes is not confined to the open ocean as nearly 70% of all hurricanes that occur over the basin reach a lifetime maximum intensity west of 60W longitude. The largest intensity increases are occurring over the Gulf of Mexico and the Caribbean Sea, where ocean temperatures are warmest and hurricanes are strongest. Decreases in intensity are noted along most of the United States coastline consistent with a hypothesis that continental aerosols act to decrease hurricane intensity.

Changes in North Atlantic Basin Hurricanes

✦Robert B. Lund, Clemson University, Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-0975, *lund@clemson.edu*

Key Words: CUSUM Statistic, Poisson Processes, Tropical Cyclones, Wind Speeds, Climate Change, Brownian Bridge

This talks overviews changes in the North Atlantic Basin tropical cyclone record since 1851. A changepoint test is first developed for categorical data via maximums of statistics that have marginal chi-squared distributions. Statistically, the asymptotic distribution of the test statistic is shown to involve the supremum of sums of squares of scaled Brownian bridge processes. The results are used to identify times of change (lack of homogeneity) in the North Atlantic Basin tropical cyclone record. We find an increase in storm counts in the 1930s that is generally attributed to the onset of aircraft surveillance. We also find a very significant increase in counts circa 1995 that does not seem to be linked to any changes in data collection techniques. The good news is that no recent changes in the individual storm strengths are found. The end conclusions are essentially the opposite of recent testimonial given in the United States Senate.

Hurricanes and Global Warming: Expectations Versus Observations

✦Christopher William Landsea, NOAA/NWS/National Hurricane Center, 11691 SW 17th Street, Miami, FL 33165, *Chris.Landsea@noaa.gov*

Key Words: climate change, hurricanes

Climate variability and any resulting change in the characteristics of hurricanes have become topics of great interest and research within the past few years. Some recent scientific articles have reported large increases in hurricane energy, numbers, and windspeeds in many basins during the last few decades in association with warmer sea surface temperatures. These increases in hurricane activity have been linked to man-made greenhouse gas changes. It is likely that some increase in hurricane peak windspeed has occurred and will continue to occur if the climate continues to warm. However, whether greenhouse gas warming is related in increases in hurricane activity is NOT the most relevant question. One needs to address instead: What is the SENSITIVITY of hurricane intensity, frequency and overall activity to greenhouse gas forcing? Is it indeed large today, or is it likely to be a small factor even several decades from now? These questions as well as an attempt to reconcile theoretial/numerical modeling studies with some recent observational papers will be addressed in the talk.

553 Emerging Statistical Problems in Finance ●

IMS, International Chinese Statistical Association

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Time-Varying Risk Premium in Large Cross-Sectional Equity Data Sets

✦Olivier Scaillet, University of Geneva and Swiss Finance Institute, , *Olivier.scaillet@unige.ch*; Patrick Gagliardini, University of Lugano and Swiss Finance Institute; Elisa Ossola, University of Lugano

Key Words: large panel, asset pricing, risk premium, thresholding, equity data, testing

We develop an econometric methodology to infer the path of the risk premium from large unbalanced panel of individual stock returns. We estimate the time-varying risk premium implied by conditional linear asset pricing models through simple two-pass cross-sectional regressions, and show consistency and asymptotic normality under increasing cross-sectional and time series dimensions. We address consistent estimation of the asymptotic variance, and testing for the asset pricing restrictions. Our approach also delivers inference for time-varying cost of equity. The empirical illustration on over 12,500 US stock returns from January 1960 to December 2009 shows that the value risk premium is counter-cyclical. It is also larger and more volatile in recent years. The cost of equity exhibits large positive and negative strays from standard unconditional estimates.

Between Data Cleaning and Inference: Pre-Averaging and Robust Estimators of the Efficient Price

✦Per Mykland, University of Chicago, Chicago, IL 60637 USA, *mykland@pascal.uchicago.edu*; Lan Zhang, University of Illinois at Chicago

Key Words: finance, high frequency data, econometrics, volatility, contiguity, large data sets

Pre-averaging is increasingly used as a method to reduce noise in financial data. So far, theory has been derived on an estimator-by-estimator basis. This paper provides a general theory for this methodology. We find a process characterization for the estimated efficient price, and inference can be done directly on the new process. This permits the study of substantially harder inference situations. For baseline volatility inference, it also shows that pre-averaging yields estimators that are arbitrarily close to being statistically efficient.

Modeling Dependence Structure via Spatial Parameter and Variable Selection

✦Chunrong Ai , University of Florida, Department of Economics, University of Florida, Gainesville, FL 32611 USA, *chunrongai@hotmail.com*; Meixia Meng, Shanghai University of Finance and Economics

Key Words: Dependence structure, Variable Selection, Spatial Parameter, Asymptotic Properties

Modeling financial data dependence properly is critical for risk management. However the existence of large number of financial instruments makes such endeavor difficult. There are three ways to overcome this difficulty. First is to specify a factor model. Second is to specify a Copulas model. Third is not to restrict dependence coefficients and let data decide via a variable selection technique. None of these approaches exploit researcher's prior knowledge about the dependence structure. For example, we all know that stock prices move more closely within sectors than between sectors. This kind of prior knowledge is not incorporated by the previous approaches. Spatial modeling on the other hand permits such knowledge and moreover incorporates other factors such as macroeconomic factors into dependence structure. In this paper, we shall model dependence structure with spatial parameter. But unlike the standard spatial models, we shall not restrict the number of spatial coefficients and instead allow data to determine via a variable selection technique. We shall derive the asymptotic properties of the proposed estimation and illustrate the procedure with an example.

554 New Approaches to Computation in Early Statistics ■

Section on Statistical Education, Section on Statistical Graphics, Section on Statistical Computing

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Using Graphics to Teach Statistics

✦Hadley Wickham, Rice University, 77004 USA, *hadley@rice.edu*

Key Words: teaching, visualisation, r

Visualisation is a powerful tool for revealing patterns in data, for communicating results, and for enticing students to study statistics. In this talk I'll discuss my experiences teaching a mixed class of graduates and undergraduates visualisation tools for exploring large datasets. Most students come in with little programming experience and little practical data analysis (although they may have had a great deal of theoretical training, which is not necessarily a help), but by the end of the class are experienced data analysts, having tackled 3-4 projects exploring datasets of at least 10,000 observations. I'll discuss why I think visualisation is such a great tool for data analysis, why using R isn't as intimidating as you might think, and how to grade data analysis projects to support the important skills of curiosity and scepticism.

Ignoring the Spatial Context in Intro Stats Classes and Some Simple Graphical Remedies

✦Juergen Symanzik, Utah State University, Department of Mathematics & Statistics, 3900 Old Main Hill, Logan, UT 84322-3900 USA, *symanzik@math.usu.edu*; Nathan D. Voge, Utah State University

Key Words: Introductory Statistics, Teaching, Maps, Graphics, Spatial Dependence, Spatial Association

Statistical data often have a spatial (geographic) context, be it countries of the world, states in the US, counties within a state, cities across the globe, or locations where measurements have been taken. However, most introductory statistics books don't hint that such data often are not independent from location, but rather are effected by some spatial association. Remedies are simple: Display data via various map views and briefly discuss additional information that can be extracted from such a graphical representation. In this talk, we will visit a variety of popular introductory statistics textbooks and show how some of the data used in examples and exercises can be initially displayed via various map views, such as choropleth maps or micromaps. We will also demonstrate how students familiar with R can create a variety of map displays via several R packages.

Gaining Inzight

✦Chris J. Wild, University of Auckland, Department of Statistics, Private Bag 92019, Auckland, 1142 New Zealand, *c.wild@auckland.ac.nz*; Dineika Chandrananda, University of Auckland

Key Words: statistical graphics, graphical user interface, statistical education, R

iNZight (“insight” with an NZ, New Zealand, pun) is gui-driven system for statistical data exploration written in R, though the user never has to interact with R. Our desire has been to provide a tool that will actively encourage the exploration of multivariate data sets and enable emphasis to be kept almost entirely on seeing what data is saying rather than learning how to drive software. We also want to keep the data students are working with always “in their faces” to minimise abstraction. We have done these things using a drag and drop metaphor in which the software is driven by dragging the names of variables from the top of the data spreadsheet and dropping them in a small number of appropriate places. What is delivered instantly, and with almost no learning curve, is graphics (plots). To obtain numerical summaries and inferential information the user has explicitly to ask for them. Plots involving up to three variables require almost no system knowledge. Relationships involving up to six variables can be explored using only very basic plot types. We discuss pedagogical imperatives and describe system capabilities, design choices and the reasoning that led to them.

555 Beyond Pharmacokinetics: Recent Advances in Science and Methodology ■

WNAR, ENAR, International Indian Statistical Association, Section for Statistical Programmers and Analysts

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Modeling PK/PGx Data

✦Donatello Telesca, UCLA Biostatistics, 650 Charles Young Boulevard S., Los Angeles, CA 90095-1772 USA, *dtelesca@ucla.edu*; Gary Rosner , Johns Hopkins; Peter Mueller, MD Anderson Cancer Center

Key Words: Graphical Models, Pharmacokinetics, RJMCMC, SNPs

We describe a general framework for the exploration of the relationships between pharmacokinetic pathways and plymorphisms in genes associated with the metabolism of a compound of interest. We integrate a population pharmacokinetics model with a simple sampling model of genetic mutation via a conditional dependence prior. Significant interactions are selected allowing the parmacokinetic parameters to depend on gene sets of variable dimension. We discuss posterior inference and prediction based on MCMC simulation. Our model finds motivation in the study of a chemotherapeutic agent used in the treatment of various solid tumors.

Ex-Vivo Modeling for Heritability Assessment and Genetic Mapping in Pharmacogenomics

✦Alison Motsinger-Reif, North Carolina State University, Raleigh, NC 27695-7566, *alison.motsinger@gmail.com*; Chad Brown, North Carolina State University; Tammy Havener, University of North Carolina at Chapel Hill; Nicholas Hardison, North Carolina State University; Eric Peters, University of North Carolina at Chapel Hill; Andrew Beam, North Carolina State University; Lorri Everrit, University of North Carolina at Chapel Hill; Howard McLeod, University of North Carolina at Chapel Hill

Key Words: heritability, pharmacogenomics, nonlinear models, association mapping

The investigation of genetic factors that determine differential drug response is a key goal of pharmacogenomics (PGX), and relies on the often-untested assumption that differential response is heritable. While limitations in traditional study design often prohibit heritability (h2) estimates in PGX, new approaches may allow such estimates. We demonstrate an ex vivo model system to determine the h2 of drug-induced cell killing and performed genome-wide analysis for gene mapping. The cytotoxic effect of 28 diverse chemotherapeutic agents on lymphoblastoid cell lines (LCLs) derived from family- and population-based cohorts was investigated. We used a high throughput format to determine cytotoxicity of the drugs on LCLs and developed a new evolutionary computation approach to fit response curves for each individual. Variance components analysis determined the h2 for each drug response and a wide range of values was observed across drugs. Genome-wide analysis was performed using new analytical approaches. These results lay the groundwork for future studies to uncover genes influencing chemotherapeutic response and demonstrate a new computational framework for performing such analysis.

Nonparametric Bayes Functional Regression for A PK/PD Semimechanistic Model

✦Michele Guindani, UT MD Anderson Cancer Center, Department of Biostatistics - Unit 1411, P. O. Box 301402, Houston, TX 77025, *mguindani@mdanderson.org*; Peter Mueller, MD Anderson Cancer Center; Gary Rosner , Johns Hopkins

Key Words: Bayesian Nonparametrics, Dirichlet Process, PK/PD models

In Clinical pharmacology, a major goal is to study the quantitative prediction of drug effects. Complex pharmacokinetic-pharmacodynamic (PK/PD) models with feedback and transition effects have been recently developed, for example to estimate the time course of myelosupression. These models typically involve the presence of covariate dependent parameters. We propose a coherent NP Bayes probabilistic framework for the analysis of these models. Our goal is to use the information from the individuals’ time courses and covariates to provide a clustering of patients according to their PK/PD profiles. This information could then be used to predict the PD time course for a patient on the basis of the observed PK profile, as well as deal with missing data.

556 Re-sampling Methods for Variance Estimation in Surveys of Governments ■

Section on Government Statistics, Section on Teaching of Statistics in the Health Sciences, Social Statistics Section, SSC, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Decision-Based Estimation for Government Statistics

✦Jun Shao, University of Wisconsin, 1300 University Ave, Department of Statistics, Madison, WI 53706 USA, *shao@stat.wisc.edu*; Sheng Wang, University of Wisconsin

Key Words: regression estimators, stratified sampling, asymptotic distribution, variance estimation

In the Annual Survey of Public Employment and Payroll and the Annual Finance Survey, a new two-stage stratified sampling design is used to reduce cost and improve data quality. The new sampling design divides each stratum into two sub-strata according to unit size and reduces the sample size of the sub-stratum with small size units. To estimate the population total, a decision-based regression estimation method was proposed in Cheng, Slud and Hogue (2010), which applies hypothesis testing to decide whether a common regression should be applied in two sub-strata. In this presentation we discuss asymptotic properties of the decision-based method as well as variance estimation for decision based regression estimators. Some empirical results are also presented.

Resampling Variance Estimation Methods for Decision-Based Stratified Regression Estimates

✦Eric Victor Slud, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *Eric.V.Slud@census.gov*

Key Words: Bootstrap, Survey regression, Estimation after testing, Collapsing Strata

The new design of the Annual Survey of Public Employment and Payroll involves PPS sampling drawn from substrata defined by small and large unit size within strata defined by state and governmental type. Survey estimates of population totals (Cheng, Slud and Hogue 2010) are assembled from stratum and substratum regression estimates after applying a hypothesis test for equality of regression slopes in deciding whether to collapse substrata. This talk describes large-sample theory and moderate-sample implementation for the bootstrap estimation of variance of the resulting survey estimates, under superpopulation models for which regression slopes within substrata are either precisely the same or slightly different.

Techniques for High Accuracy in Estimation of the Mean Squared Prediction Error in General Small-Area Model

✦Ansu Chatterjee, University of Minnesota, School of Statistics, 224 Church Street SE, Minneapolis, MN 55455 USA, *chatterjee@stat.umn.edu*; Partha Lahiri, University of Maryland at College Park

Key Words: Small domain, resampling methods, sample survey, mixed models, risk function, EBP

A general small area model is a hierarchical two-stage model, of which special cases are mixed linear models, generalized linear mixed models and hierarchical generalized linear models. In such models, the variability of predictors (like the empirical best predictor or the empirical best linear unbiased predictor)can be quantified with their mean squared prediction error (MSPE), or other risk functions. Estimators for the MSPE are generally not available outside some special cases. First, we propose a simple resampling-based estimation of MSPE for any general small area model. Second, we propose three techniques for improving on the basic, resampling-based MSPE estimator to achieve high order accuracy. Computational issues and other properties of these improved MSPE estimators will be discussed.

557 Statistical Analysis of Populations of Images ■●

JASA, Applications and Case Studies, International Indian Statistical Association, Section on Health Policy Statistics, Section on Statistical Computing, Section on Statistical Graphics

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Population Value Decomposition: A Framework for the Analysis of Image Populations

✦Ciprian Crainiceanu, Johns Hopkins University, 615 N Wolfe Street, Baltimore, MD 21205, *ccrainic@jhsph.edu*

Key Words: signal extraction, EEG, fMRI, statistical modeling

Images, often stored in multidimensional arrays are fast becoming ubiquitous in medical and public health research. Analyzing populations of images is a statistical problem that raises a host of daunting challenges. The most severe challenge is that data sets incorporating images recorded for hundreds or thousands of subjects at multiple visits are massive. We introduce the population value decomposition (PVD), a general method for simultaneous dimensionality reduction of large populations of massive images. We show how PVD can seamlessly be incorporated into statistical modeling and lead to a new, transparent and fast inferential framework. Our methodology was motivated by and applied to the Sleep Heart Health Study, the largest community-based cohort study of sleep containing more than 85 billion observations on thousands of subjects at two visits.

558 Innovative Strategies for Incorporating Consulting into Graduate Education

Section on Statistical Consulting, Section on Quality and Productivity, Section on Teaching of Statistics in the Health Sciences

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Educating Statisticians to Become Interdisciplinary Collaborators

✦Eric Vance, Virginia Tech, 212 Hutcheson Hall (0439), Blacksburg, VA 24061, *ervance@vt.edu*

Key Words: statistical consulting, interdisciplinary collaboration, video-recorded meetings, LISA, student education, student training

Graduate students in the Laboratory for Interdisciplinary Statistical Analysis (LISA) at Virginia Tech are trained to become effective statistical collaborators. They learn how to manage, display, analyze, and write about data in an introductory course and then learn and apply effective communication skills in the course "Communication in Statistical

Collaborations," which aims to teach them what they need to know to be effective statistical collaborators that they don't learn in their more technical statistics courses. In their third and future semesters students work in LISA: 1) Teaching short courses on statistical topics for graduate students from other departments to apply in their research; 2) Answering quick statistical questions in LISA Walk-in Consulting; 3) Providing advice on experimental design, modeling, data analysis, and interpretation of statistical results to researchers across campus; 4) Collaborating on multidisciplinary research projects around the world. Students continue their training by giving and receiving feedback and by analyzing video-recorded meetings to learn and practice strategies for improving their statistical consulting and collaboration skills.

Iteratively Growing Our Web

✦Xiao Chen, UCLA ATS/Statistical Consulting Group, 5308 Math Sciences, Los Angeles, CA 90095 USA, *xiao.chen@ucla.edu*

Key Words: web, interaction, graduate students, stat software

The UCLA ATS Statistical Consulting Group has used its website to document and share knowledge in statistical computing for more than a decade. It started as an experiment to help more efficiently UCLA researchers and graduate students with their statistical computing needs and the website has proven to be extremely successful. It has become well-known not only to the UCLA research community but also to the statistical computing community worldwide. We believe that the success of our website comes largely from the interactions of the group with many UCLA graduate students and researchers and from our constant learning of new statistical computing techniques. The growing of our website has been an iterative process. In this paper, we will focus on describing this iterative process with a couple of examples. We will also discuss our experiments with different ways for delivering our material online, including podcasts, social bookmarks, and streaming video. Last, we will discuss the challenges we face to keep the site accurate and up to date.

Innovative Strategies for Incorporating Consulting into Graduate Education

✦Nagaraj K. Neerchal, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, *nagaraj@math.umbc.edu*; Matthias Gobbert, UMBC

Key Words: Consulting, Interdisciplinary, professional development

UMBC, a small public university of 9,000 undergraduate and 3000 graduate students, has gained recognition for significant growth in interdisciplinary research in the last decade and a half. The Center for Interdisciplinary Research and Consulting (CIRC) is a joint venture between the Applied Mathematics and Statistics graduate programs at UMBC. A unique feature of CIRC is that it promotes the idea of graduate students themselves forming teams of fellow graduate students from other departments and proposing interdisciplinary projects and lining up faculty help. This has resulted in publications with graduate students from different departments as co-authors. This experience in business development has not only created attractive resumÈ entries but also has helped the graduate students develop contacts on and off campus. CIRC has formed a strong alliance with the campus initiative of creating a High Performance Computing Facility for training students. This collaboration has resulted in a NSF award for a Research Experience for Undergraduates (REU) site on High Performance Computing for 2010 and 2011.

Tales from Mozambique as an 'On-the-Ground' Statistician

✦Mark Thomas Seiss, Laboratory for Interdisciplinary Statistical Analysis, 621 Progress St, Blacksburg, VA 24060, *maseiss@gmail.com*

Key Words: Statistical Collaboration, Survey, Water, Mozambique

As a statistics Ph.D. student and LISA (Laboratory for Interdisciplinary Statistical Analysis) statistical collaborator at Virginia Tech, the opportunity arose to collaborate on a multidisciplinary research project in Mozambique to assess the effectiveness of drilling borewells in rural villages without a sufficient supply of clean water. I will first discuss my efforts to develop the overall research questions in collaboration with the other partners on the project in the months leading up to the field work. This includes the design of specific survey questions needed to properly answer these research questions. I will then speak about my experience embedded with the other researchers in Mozambique doing nightly data checks and analyses to ensure the quality of the data. Any issues that may arise and techniques used to address them will also be discussed. I will conclude my talk by speaking about some of the statistical models we have developed to answer the research questions and the role of this unique collaborative experience in my training to be an effective statistical collaborator. This talk will include photos and short segments of video to illustrate my experiences.

559 Memorial Session for Nathan Keyfitz and Zakkula Govindarajulu

Memorial

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

The Survey Sampling Contributions Of Nathan Keyfitz

✦Gary Shapiro, Statistics Without Border, , *g.shapiro4@verizon.net*

Nathan Keyfitz made important contributions in many areas of mathematics, demography, and statistics. This paper concentrates on two major advances to the practice of survey sampling that Keyfitz made early in his career. One deals with the redesign of a survey in which it is desired to retain as many first stage sampling units as possible from the old design in the new design. This is subject, however, to not changing the probability of selection over all possible samples for any first stage unit. The second advance by Keyfitz was the development of a new method of variance estimation for sample surveys. Keyfitz is the original developer of the linearization method for estimating variances, which is one of the two primary variance estimation methods used today, more than 50 years after Keyfitz's ground-breaking paper.

Simultaneous Testing And Estimation In Linear Models

✦Hossein Mansouri, Texas Tech University,

Simultaneous tests about linear functions of parameters of a linear model are constructed to control the familywise error rate at a prescribed level. Likewise, simultaneous confidence intervals are constructed to control the overall confidence level. In this presentation, the strengths and weaknesses of several well-know techniques for simultaneous inference will be discussed and a study of the small sample properties of these techniques under different distributional assumptions will be presented. We will also study the stability of the sampling distributions of the test statistics for the least square technique as compared to some rank-based methods for multiple comparisons of the parameters of a linear model.

Sequential Confidence Limits For The Ratio Of Binomial Proportions Ii: Unequal Sample Sizes

✦Hokwon Cho, University of Nevada,

In this paper, we consider the approximate sequential confidence limits for the ratio of two independent binomial proportions that have unequal sample sizes. By defining a sample-ratio, we generalize results of Cho and Govindarajulu (2008) for binomial variates with equal sample sizes. Due to the nonexistance of an unbiased estimator for the ratio, we develop the procedure based on a modified maximum likelihood estimator. We also investigate the large-sample properties of the proposed estimator and its finite sample behavior through numerical studies.

Buckley-James Estimator And Its Variations

✦Mai Zhou, University of Kentucky, 40506, *mai@ms.uky.edu*

The Buckley-James estimator (1979) is the generalization of the least squares estimator for linear regression models when the responses are subject to right censoring (AFT model). We discuss an alternative definition of the Buckley-James estimator,and study the similarity/differences. We highlight some benefits this new definition brings. Examples will be given.

Misclassication Probabilities For N-Fold Random Samples

✦Michael Klass, University of California, Berkeley,

Suppose it is of interest to classify a new distribution on the basis of a random sample of size n as being either P0 or P1, the support of which is a finite set f1; 2; : : : ;Lg. Nothing is known about the distributions P0 and P1 except that we have training samples of size n0 and n1 from P0 and P1, respectively. Assuming that n0 an and n1 bn for some a; b > 0 it is shown that the probability of misclassification goes to zero exponentially as n becomes large. Our method of classification employs the data from all three samples to construct two distinct pairs of maximum likelihood distributions for P0 and P1, the first assuming the sample to be classified has distribution P0, the second assuming it has distribution P1, deciding how to classify the sample in question by means of a likelihood ratio statistic.

560 In Honor of Jay Kadane's 70th Birthday: Decision Theory in all Walks of Life ■●

WNAR, Section on Bayesian Statistical Science

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

In Honor of Jay Kadane's 70th Birthday: Decision Theory in All Walks of Life

✦Dalene Stangl, Duke University, 27708, *dalane@stat.duke.edu*; ✦Bob Winkler, Duke University, 27708, *rwinkler@duke.edu*; ✦Giovanni Parmigiani, Harvard/Dana-Farber, 02115, *gp@jimmy.harvard.edu*; ✦Don Berry, MD Anderson, 77030, *dberry@mdanderson.org*; ✦Teddy Seidenfeld, Carnegie Mellon University, 15213, *teddy@stat.cmu.edu*; ✦Mark Schervish, Carnegie Mellon University, 15213, *mark@stat.cmu.edu*

Key Words: Bayesian, Kadane

Joseph B. (Jay) Kadane is well-known for his contributions to Bayesian theory, econometrics, and a wide variety of fields of application. Jay is a cross-disciplinary statistician, having worked in econometrics, law, medicine, political science, sociology, computer science, archaeology, and environmental science, among others. His work truly embodies the theme of this year's JSM, "Statistics: An All-encompassing Discipline". In celebration of Jay's 70th birthday, several colleagues and former students will highlight his contributions, which run the gamut from the theoretical foundations of statistics to the practicalities of testifying as an expert witness in a racial discrimination case of speeding on the New Jersey Turnpike.

561 Innovative Analysis of Clinical Trial Data in US VA Cooperative Studies Program ■●

Section on Government Statistics, Section on Health Policy Statistics, Social Statistics Section, WNAR, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Multivariate Survival Model Using Latent Variables With The Department Of Veterans Affairs Diabetes Trial (Vadt) Data

✦Gideon D Bahn, Hines Edward VA Hospital CSPCC, 5000 5th Ave., Hines, IL 60141, *gideon.bahn@va.gov*

Key Words: Multivariate, Survival Model, Latent Variables, Veterans Affairs Diabetes Trial (VADT), Structural Equation Modeling

Framework: Conventionally, the Cox survival model has one outcome as a single and non-recurring discrete variable, though often several related events are made into one composite outcome. This model has limitation; losing important measured outcomes by not counting recurring events within the same person. This problem has been dealt with in frailty models. This study, however, shows flexibility of a survival model using a latent variable (SML), measuring recurring events. The frailty models are still limited to analyzing either multiple but different events or the reoccurrence of the same event. This study utilizes the continuous time survival model in a Structural Equation Modeling (SEM) framework. First, Cox survival model and a SML are compared. Second, frailty models and SML with recurring events are compared. Then, a multivariate survival model using multiple latent variables (MSML) is investigated using SEM. The results show the similarity between the first two comparisons, but the MSML provides additional information that is different from the first two. Moreover, the MSML may be explored compared with multi-stage and/or multivariate frailty models. The Department of Vete

Study Medication Compliance And Its Impact On Finding Treatment Efficacy

✦Kousick Biswas, VA Cooperative Studies Program, VA Maryland Healthcare System, PO Box 1010, Perry Point, MD 21902, *Kousick.Biswas@va.gov*

Key Words: Clinical Trial, Compliance, Efficacy, Experimental Treatment

The efficacy of any experimental treatment depends on the treatment-compliance by the participants. The treatment-compliance in a clinical trial setting is measured using two most commonly used quantities - self reported use and pill count. Both of these measures are unreliable as these can be manipulated by the participants to report a higher compliance. In the experimental treatment research for substance abuse, a series of recently completed clinical trials ended being negative even though the experimental treatments were found to be beneficial in smaller pilot studies. In these studies compliance was measured using self report and pill counts and reported in the 85% - 95% range. In a recently completed negative study, an ad-hoc treatment-compliance analysis was undertaken using treatment concentration in the urine and a very little agreement was reported when compared with usual compliance measures. The negative findings in the efficacy outcomes may be attributed to a very low actual treatment-compliance. This finding calls for a serious discussion about a well known problem in clinical trials for experimental treatments for non-life threatening indications - measurement of tre

Subgroup Analysis: Guidelines For Department Of Veterans Affairs Clinical Trials

✦Robert Lew, VA Boston Coop. Studies Prog. Coordinating Ctr., 150 S. Huntington Avenue, Boston, MA U.S.A., *Robert.Lew2@va.gov*; Ying Lu, Palo Alto VA CSP Coordinating Center and Stanford ; Mei-Chiung Shih, Veterans Administration; Grant Huang, Veterans Administration

Key Words: subgroup analysis, type I error, policy guidelines

The VA has done over 500 multicenter clinical trials drawing most subjects from the millions of US veterans cared for at VA facilities. The typical trial primary hypothesis has 90% power and a 5% Type I error. Subgroup analyses are exploratory because the sample size adequate to power all of them would make the trial infeasible. Thus, what is a fair level of Type I error for subgroup analyses? We propose guidelines that would: 1) pre-specify a list of subgroup analyses, 2) require analytic reports to reference the primary and other pertinent results based on the primary data, and 3) set the level below 5% but above the Bonferroni level. The pre-specified list would include hypotheses in the proposal and those recommended by the DMC during the trial and by disinterested experts. We discuss approaches to Type I error that account for correlation among the test statistics associated with a set of hypotheses; Efron's method to control for the false discovery rate, Simon's empirical Bayes method for clinical trials, and Bayesian subgroup analysis methods. Adherence to such guidelines would refine secondary analyses, inter-relate conclusions, and clarify the import of major results.

An Approach To Analyzing Multiple Correlated Adverse Events

✦Zhibao Mi, VA Cooperative Studies Program, CSPCC 151E, Research Bldg. 4, Boiler House Rd, PO BOX 1010, Perry Point, 21902 US, *Zhibao.mi@va.gov*; Dengfang Tang, VA Cooperative Studies Program; Xiaoli Lu, VA Cooperative Studies Program; Kousick Biswas, VA Cooperative Studies Program; Joseph Collins , VA Cooperative Studies Program

Key Words: Adverse Effect, Multiple correlated events, Rate ratio, Hazard ratio, Competing risks

Adverse effect assessment involves many clinical symptoms, and each individual may experience multiple adverse events (AE). A common practice for AE analysis that compares proportions of AE or the times to first event is simple and easy to implement, but it may lose data information. We propose using rate ratio (RR) calculate from AE incidence rates (IR) and hazard ratio (HR) to assess adverse effects for multiple correlated AE data. For analyzing event counts, AE specific IRs and the corresponding RRs are calculated and tested for statistical significance by computing exact probabilities, and the significance level is adjusted for multiple AE strata. The overall RR is tested using a counting process method. For analyzing times to multiple events, AE specific HRs are tested using the Cox model and overall HR is estimated and tested based on the Anderson-Gill and WLW marginal competing risk models. Multiple correlated adverse event data from a VA clinical trial was used as an example. From the analysis, HR estimated from the Anderson-Gill model was close to the overall RR. When compared with the frailty and conditional risk set models, the marginal model yielded similar HR.

Evaluation Of Treatment Effect On Gap Time

Bruce Chow, US VA Cooperative Studies Program Palo Alto Coordinating Center; ✦Julia Y Lin, US VA Cooperative Studies Program Palo Alto Coordinating Center, 701B Norther Shoreline Blvd., Mountain View, CA 94043, *julia.lin@va.gov*; Mei-Chiung Shih, Veterans Administration; Ying Lu, Palo Alto VA CSP Coordinating Center and Stanford

Key Words: Gap Time, Clinical Trials, Survival Analysis, Causal Inference, Informative Censoring, Selection Bias

Times of sequential event are observed in some clinical trials, and the time from the first to the second event is the gap time. In a randomized control trial of smoking cessation treatments, 943 smokers with military-related PTSD were randomized to smoking cessation treatment integrated with mental health care (integrated care; IC) or the usual care of referral to smoking cessation clinics (SCC), and they were followed up for 18 to 48 months (McFall et al. 2010). We are interested in estimating the causal effect of the IC on time to quitting smoking, and then the causal effect of the IC on the gap time of quitting to relapse, compared to the SCC. The treatment effect on relapse suffers from informative censoring and selection bias due to its conditional nature on the occurrence of quitting. We utilize a principal stratification framework (Frangakis & Rubin 1999, 2002) that precludes the confounding when stratifying on observed post-randomization behavior (i.e., observed quitting).

562 Ranked Set Sampling ●

Section on Nonparametric Statistics

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Judgment Post-Stratification For Estimating Success Probability With Nonresponse

✦Tian Chen, The Ohio State University, 2825 Neil Ave apt818, columbus, OH 43202, *chen.1476@buckeyemail.osu.edu*; Elizabeth Stasny, Ohio State University

Key Words: judgment post-stratification, nonresponse

We compared judgment post-stratification with simple random sample methods for estimating the probability of success when analyzing binary data subject to nonresponse. We believe that when nonresponse is related to the variable of interest, using judgment post-stratification provides additional information and leads to a reduction in bias under both perfect ranking and imperfect ranking. We illustrate this using primary care physician data on the use of Health Information Technology from the 2010 Ohio Medicaid HIT Survey for Health Care Professionals.

Quantile Inference Based On Partially Rank Ordered Set Samples

✦Omer Ozturk, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *omer@stat.osu.edu*

Key Words: Imperfect ranking, ranking models, median confidence interval, partially ranked set, sign test, judgment ranking

In this talk, we develop statistical inference for population quantiles based on partially rank ordered set sample (PROSS) designs. The PROSS sample is similar to a ranked set sample with some clear differences. This design first creates a partially rank ordered subsets by allowing ties whenever the units in a set can not be ranked with high confidence. It then selects a unit for full measurement at random from one of these partially rank ordered subsets. Since ranking process in PROSS design utilizes full potential of rankers, it usually has a smaller ranking error than a ranked set sample of the same size. The paper develops a point estimator, confidence interval and hypothesis testing procedure for the population quantile of order p. Exact as well as asymptotic distribution of the test statistic is derived. It is shown that the null distribution is distribution free and statistical inference is reasonably robust against any possible ranking error in ranking process.

Estimation Of The Distribution In Selection Bias Models: A Ranked Set Sampling Approach

✦Kaushik Ghosh, University of Nevada, Las Vegas, , *kaushik.ghosh@unlv.edu*; Ram C Tiwari, FDA/CDER/OTS/OB

Key Words: ranked set sampling, biased sampling, empirical distribution

In this talk, we will discuss the problem of estimation of the underlying distribution when observations are collected as ranked set samples from various weight-biased versions of this distribution. We propose a nonparametric estimator of the distribution and investigate its large-sample properties. Finally, we look at the effect of imperfect ranking and present the results of a simulation study.

Using Ranked Set Sampling With Hierarchical Linear Models

✦Xinlei (Sherry) Wang, Southern Methodist University, 3225 Daniel Avenue, Dallas, TX 75275 USA, *swang@smu.edu*; Lynne Stokes, Southern Methodist University; Holly Stovall, Southern Methodist University

Key Words: hierarchical linear models, judgement post-stratification, order statistics, power analysis, ranking

Ranked set sampling (RSS) is a well established method of data collection, which is useful in situations where the variable of interest is expensive to measure, but sampling units can be easily recruited and ranked by some means not requiring quantification. It utilizes the assigned ranks to provide auxiliary information about the measured units. It has been proved to provide improved estimators of the mean, variance and distribution functions over simple random sampling (SRS) of the same size. In this work, we will consider ranked set sampling for use with hierarchical linear models (HLM) that have a wide range of applications in education sciences, to achieve potential improvement in the parameter estimation as well as the statistical power . We will consider ranking based on covariates at the school level and ranking based on covariates at the student level. In addition, we consider judgement post-stratification, a variant of RSS that has an underlying SRS if ignoring the rank information, for use with HLMs. Numerical results from simulation and empirical studies will be provided to examine the performance of our proposed approach.

Design-Based Estimation for Ranked-Set Sampling in Finite Populations

✦Mohammad Jafari Jozani, University of Manitoba, Department of Statistics, 325 Machray Hall, Department of Statistics, Winnipeg, MB R3T 2N2 Canada, *m_jafari_jozani@umanitoba.ca*; Brad Johnson, University of Manitoba, Department of Statistics

Key Words: Finite Population, Hansen-Hurwitz Estimator, Horvitz-Thompson Estimator, Inclusion Probability, Ranked Set Sampling

In this talk, we consider design-based estimation using ranked set sampling (RSS) in finite populations. We first derive the first and second-order inclusion probabilities for an RSS design and present two Horvitz-Thompson type estimators using these inclusion probabilities. We also develop an alternate Hansen-Hurwitz type estimator and investigate its properties. In particular, we show that this alternate estimator always outperforms the usual Hansen-Hurwitz type estimator in the simple random sampling with replacement design with comparable sample size. The theoretical results are augmented by numerical and simulation studies as well as a case study using a well known data

set. These show that RSS design can yield a substantial improvement in efficiency over the usual simple random sampling design in finite populations.

563 Section on Statistical Computing Student Award Winners ■

Section on Statistical Computing, Section for Statistical Programmers and Analysts, Section on Statistical Graphics

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Community Extraction For Social Networks

✦Yunpeng Zhao, Department of Statistics, University of Michigan, 439 West Hall, 1085 South University Ave, Ann Arbor, MI 48109, *yunpeng@umich.edu*; Liza Levina, University of Michigan; Ji Zhu, University of Michigan

Key Words: social networks, community detection, block model

Analysis of networks and in particular discovering communities within networks has been a focus of recent work in several fields and has diverse applications. Most community detection methods focus on partitioning the entire network into communities, with the expectation of many ties within communities and few ties between. However, many networks contain nodes that do not fit in with any of the communities, and forcing every node into a community can distort results. Here we propose a new framework that focuses on community extraction instead of partition, which allows for weakly connected nodes. The main idea is that the strength of a community should not depend on ties between members of other communities, only on ties within itself and ties to the outside world. The proposed extraction criterion has a natural probabilistic interpretation in a wide class of models and performs well on simulated and real networks. Under the additional assumption of the block model, we establish asymptotic consistency of estimated node labels and propose a hypothesis test for determining the number of communities.

Simultaneous Multiple Response Regression And Inverse Covariance Matrix Estimation Via Penalized Gaussian Maximum Likelihood

✦Wonyul Lee, University of North Carolina at Chapel Hill, Department of Statistics and Operations Research, CB3260, University of North Carolina, Chapel Hill, NC 27599, *wonyull@email.unc.edu*; Yufeng Liu, Department of Statistics and Operation Research, UNC-Chapel Hill

Key Words: Covariance estimation, Joint estimation, LASSO, Multiple response, Regression, Sparsity

Many multivariate regression techniques are designed for univariate response cases. For problems with multiple response variables available, one common approach is to apply the univariate response regression technique separately on each response variable. Although it is simple and popular, the univariate response approach ignores the information among response variables. In this paper, we propose two new methods for utilizing joint information among response variables. Both methods are in a penalized likelihood framework with weighted L1 regularization. The proposed methods provide sparse estimators of conditional inverse covariance matrix of response vector given explanatory variables as well as sparse estimators of regression parameters. Our first approach is to estimate the regression coefficients with plug-in estimated covariance matrices, and our second approach is to estimate the regression coefficients and the covariance matrix simultaneously. Asymptotic properties of our methods are explored. Through several simulated examples and application to a real Glioblastoma cancer data set, we demonstrate that the proposed methods perform competitively.

Visual Statistical Inference For Regression Parameters

✦Mahbubul Amin Majumder, Iowa State University, 1418 Snedecor Hall, Ames, IA 50010 USA, *mahbub72@gmail.com*; Heike Hofmann, Iowa State University; Dianne Cook, Iowa State University

Key Words: Statistical graphics, Visual inference

Statistical graphics play a crucial role in exploratory data analysis, model checking and diagnosis. Until recently there were no formal visual methods in place for determining statistical significance of findings. This changed, when Buja et al.(2009) conceptually introduced two protocols for formal tests of visual findings. In this paper we take this a step further by comparing the lineup protocol (Buja et al.2009) against classical statistical testing of the significance of regression model parameters. A human subjects experiment is conducted using simulated data to provide controlled conditions. Results suggest that the lineup protocol provides results equivalent to the uniformly most powerful (UMP) test and for some scenarios yields better power than the UMP test.

564 Statistical Analyses of Judging in Athletic Competitions: The Role of Human Nature ■●

Section on Statistics in Sports, Social Statistics Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

The Phantom Judge

✦John Emerson, Yale University, P.O. Box 208290, New Haven, CT 06520-8290 USA, *john.emerson@yale.edu*

Key Words: figure skating, scoring system, bias, randomization

In figure skating, past controversy over judging has led to new rules that invoke a random selection of judges. That randomness can, and in fact has, affected the outcome of competitions. This talk explores issues surrounding this rule change, comparing it to the system used in Olympic Diving, where the reporting of scores is completely transparent.

Statistical Sleuthing By Leveraging Human Nature: A Study Of Olympic Figure Skating

✦Taylor Baillie Arnold, Yale University, 24 Hillhouse Ave, New Haven, CT 06511, *taylor.arnold@yale.edu*

Key Words: Figure Skating, goodness-of-fit, hypothesis test, Kolmogorov-Smirnov test

Analysis of figure skating scoring is notoriously difficult under the new Code of Points (CoP) scoring system, created following the judging scandal of the 2002 Olympic Winter Games. The CoP involves the selection of a random subpanel of judges; scores from other judges are reported but not used. An attempt to repeat the methods of previous studies in CoP scoring failed to recreate the competition scores from the raw scoring sheets. This raised the concern that different subpanels of judges were being selected for each skater (breaking ISU rules). However, it is also possible that the ISU was attempting to further reduce transparency in the system by permuting, separately for each skater, the order of the presentation of scores from the judging panel. Intuition suggests that it is impossible to tell the difference between accidental randomization and intentional permutation of the judges' scores. Although the recent changes do successfully prevent the study of nationalistic bias, this paper provides strong evidence against the hypothesis that a separate random subpanel is chosen for each competitor. It addresses the problem by applying Conover's extension of the K-S test.

Perception = Reality? Analyzing Specific Allegations Of Nba Referee Bias

✦Ryan M. Rodenberg, Florida State University, 139 Chieftan Way, 1006 Tully, Tallahassee, FL 32306-4280, *rrodenberg@fsu.edu*

Key Words: Bias, Referee/Judges, Basketball

The National Basketball Association's (NBA) investigation of a gambling scandal involving one of its referees put allegations of basketball referee bias in the spotlight. This paper analyzes specific allegations of bias by Miami Heat coach and general manager Pat Riley against NBA referees Steve Javie and Derrick Stafford. In the course of analyzing every referee who officiated a Miami Heat game during a nine year period, neither Javie nor Stafford exhibited any systematic bias that had an adverse effect on the Miami Heat. In fact, the Heat performed slightly better than predicted when Javie officiated their games. The results provide real-world empirical evidence consistent with "confirmation bias," a theory grounded in the finding that individuals with a vested interest in certain self-justifying outcomes may reach generalized conclusions unsupported by actual evidence.

Bias In Sequential Order Judging: Primacy, Recency, Sequential Bias, And Difficulty Bias

✦Kurt W Rotthoff, Seton Hall University, 400 South Orange, JH 621, South Orange, NJ 07079, *kurt.rotthoff@shu.edu*; Hillary Morgan, Seton Hall University

Key Words: Sequential Order Judging, Bias, Primacy, Recency, Sequential Bias, Difficulty Bias

Does order matter? Studies have found that going first or last in a sequential order contest leads to a biased outcome, commonly called primacy and recency. This study tests for those biases and two other forms: sequential bias and difficulty bias. We confirm previous work, finding that it is valuable to go later in the competition. We find no evidence of sequential bias, meaning one person's score is unrelated to the previous person's score. Finally, we find strong evidence of difficulty bias. Contestants who attempt higher difficulty increase their execution score, even when difficulty and execution scores are judged separately.

565 Calibration Weighting and Benchmarking in Sample Surveys ■

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

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Calibration Estimation And Longitudinal Survey Weights: Application To The Nsf Survey Of Doctorate Recipients

✦Michael D Larsen, George Washington University, Biostatistics Center, 6110 Executive Blvd, Suite 750, Rockville, MD 20852, *mlarsen@bsc.gwu.edu*; Siyu Qing, George Washington University; Beilei Zhou, George Washington University; Mary A Foulkes, George Washington University

Key Words: calibration weighting, longitudinal study, panel study, raking, SESTAT, survey sampling

The NSF's Survey of Doctorate Recipients is conducted every 2 or 3 years and collects detailed information on individuals receiving PhDs in science and engineering in the U.S. and some others with PhDs from abroad in these areas. Survey weights adjust for oversampling and nonresponse on a cross sectional basis. A significant portion of the sample (e.g., 60% on 3 or more surveys from 1993-2006) appears in multiple survey years and can be linked across time. No longitudinal weight exists that would enable estimation of statistical models or comparison of finite population characteristics using data from multiple survey waves together. This paper explores calibration estimation for construction of such a longitudinal weight. Three requirements are considered when producing longitudinal weights. First, the weight needs to be calculable from existing data. Second, the weight needs to be useful for reproducing key cross sectional analyses. Third, the weight should be low in variability. Choices of initial weights and calibration targets are compared in a series of analyses.

Exploring Calibration Of Nis Weights Utilizing Telephone Status From The Nhis

✦Meena Khare, National Center for Health Statistics, CDC, USA, , *mkhare@cdc.gov*; Kennon R. Copeland, NORC at the University of Chicago ; Abera Wouhib, NCHS/CDC; Nadarajasundaram Ganesh, NORC at the University of Chicago; Kirk M Wolter, NORC at the University of Chicago

Key Words: RDD telephone surveys, noncoverage adjustment

National Immunization Survey (NIS), a landline telephone survey, monitors vaccination coverage rates among children aged 19-35 months and adolescents aged 13-17 years. In 2009, the National Health Interview Survey (NHIS) attempted to obtain consent to contact vaccination providers of children with a completed interview. This NHIS Provider Record Check (NHIS-PRC) study followed the NIS provider data collection procedure for age-eligible children and adolescents with consent. The NHIS-PRC data collection operation covered both landline and non-landline telephone households. The increased number of inaccessible households in recent years due to non-landline telephone status widens the gap between the target and the sampled populations, raising a major concern regarding the representativeness of the NIS sample. We explore the impact of the potential nonrepresentativeness of the NIS sample on estimated vaccination coverage rates. Distribution of telephone status, household characteristics, and estimated vaccination coverage rates from the 2009 NHIS-PRC study are utilized to calibrate 2009 NIS estimates. We compare estimated vaccination coverage rates with and without calibrated weights.

Wtadjx Is Coming: Calibration Weighting In Sudaan When Unit Nonrespondents Are Not Missing At Random And Other Applications

✦Phil Kott, RTI International, 6110 Executive Blvd., Rockville, MD 20852 USA, *pkott@rti.org*

Key Words: Response model, Nonresponse bias, Generalized raking, Variance estimation, Optimal calibration

RTI's SUDAAN 11 will include a new weight calibration procedure, WTADJX, which will allow the set of model variables governing the response model to differ from the set of benchmark variables. As a result, the new procedure may be used to assess and perhaps even correct for unit nonresponse that is not missing at random; in other words, when unit response is a function of variables known only for the respondents. The WTADJUST procedure, introduced in SUDAAN 10, produces calibrated weights by implementing an extension of generalized raking. By forcing the weighted totals for a set of "calibration variables" to equal benchmark totals computed from either the whole sample, a larger sample, or the entire population, WTADUST-produced calibration weights can reduce or eliminate the potential for bias from unit nonresponse under a reasonable response model. With WTADJX, the response-model and the benchmark variables will be allowed to differ. Other potential uses of this procedure include producing well-behaved calibrated weights asymptotically identical to "optimal" ones and simplifying the computation of replicate weights when using WTADJUST to produce calibration weights.

Simultaneous Calibration And Nonresponse Adjustment With Explicit Variance Formulas

Eric Victor Slud, U.S. Census Bureau ; ✦Yves Thibaudeau, U.S. Census Bureau, Washington, DC 20233, *yves.thibaudeau@census.gov*

Key Words: selection probabilities, nonresponse adjustments, design-consistency, linearization formula

Statistical agencies invest considerable resources in the scientific design of survey samples, one aspect of which is the availability of explicit unit and joint selection probabilities. In textbook theory, these selection probabilities enable design-consistent estimation of the variance through linearization formulae. In practice, however, the inverse selection- probability weights are subject to multiple rounds of adjustments to reflect population benchmarks and to compensate for nonresponse. After these operations the weights no longer have a clear interpretation. Then, formula-based design-consistent variance estimation methods are often not longer easily available, and replication techniques are used instead. We propose a comprehensive optimization-based method for adjusting the initial survey weights for population benchmarking and nonresponse and also accommodating weight compression or smoothing, in a single operation. The benefit is a unified theoretically supported weight adjustment with guaranteed bounds on the weights, and which maintains design-consistency under nonresponse or linear-model assumptions.

Two Stage Bayesian Benchmarking For Small Area Estimation

✦Malay Ghosh, University of Florida, , *ghoshm@stat.ufl.edu*

Key Words: Bayes, Benchmarking, Small Area, Two Stage

Small area estimates are usually model based, and thus when aggregated, do not typically match the direct estimate for a large geographical area. This often causes concern since the direct estimate for a large geographical area is believed to be quite reliable. In order to address this concern, and possibly also to guard against potential model failure, small area estimates are often adjusted to match the direct estimate for the larger area after aggregation. This is usually referred to in the small area literature as benchmarking. The paper proposes two-stage Bayesian benchmarking with one single model. A simple illustration is where one wants to benchmark the state estimates to the national estimate, and then the county estimates to the corresponding benchmarked state estimates.

566 How can the noninferiority margin be defined where controlled trials historically lack? ■●

Biometrics Section, Biopharmaceutical Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Noninferiority Trial Designs and Issues: Community-Acquired Bacterial Pneumonia (CABP)

✦Thamban Valappil, Division of Biometrics IV, Office of Biostatistics, FDA, 10903 New Hampshire Avenue, Silver Spring, MD 20993, *thamban.valappil@fda.hhs.gov*; ✦Daniel Rubin, FDA, Office of Biostatistics, 10903 New Hampshire Avenue, Silver Spring, 20993, *Daniel.Rubin@fda.hhs.gov*

Key Words: noninferiority, Clinical trials, CABP, anti-infective

Clinical trials evaluating antibacterial drugs for community-acquired bacterial pneumonia have used non-inferiority designs, and a primary endpoint of investigator-assessment of clinical response, defined as resolution of signs and symptoms of the disease 1-2 weeks after the end of therapy. However, historical data that reliably estimate the magnitude of the active control effect over placebo are not available for this endpoint, and a non-inferiority margin cannot be rigorously justified. There is stronger evidence for a non-inferiority margin when using an all-cause mortality endpoint, but recent registration trials have enrolled patients with low risk for mortality. We will discuss design and analysis considerations in current trials, including the use of endpoints defined early in the treatment course, the impact of prior antimicrobial therapy on non-inferiority results, and concerns about the use of biomarkers that do not capture patient feeling, function, or survival.

Noninferiority Assessments in Clinicial Trials: A Nightmare in Practice ?

✦Hans Ulrich Burger, Hoffmann-La Roche, Z‰hringerstr. 58, Neuenburg, 79395 Germany, *hans_ulrich.burger@roche.com*

Key Words: Non-inferiority, Risk-benefit assessment, PFS, Overall survival

In situations were standard superiority trials for efficacy are not applicable, non-inferiority designs are an alternative to demonstrate efficacy of a new compound as well as to allow an adequate risk benefit assessment. The basic design features of such trials are known since quite some time, the implementation in practise can pose, however, still a number of challenges, in terms of design as well as in the interpretation of the results. This talk will present some examples and by that highlight the difficulties we still have today with such designs, in the definition of a suitable non-inferiority margin, in the interpretation of results and in its potential regulatory implications. Please note Sue Jane Wang is the session organizer

Evaluating Noninferiority Trials for Combination Therapies

✦Heinz Schmidli, Novartis Pharma AG, P.O.Box, Basel, CH-4002 Switzerland, *heinz.schmidli@novartis.com*

Key Words: meta-analysis, non-inferiority, clinical trial, evidence synthesis, random effect, historical information

Placebo controlled clinical trials are not always possible for ethical reasons, and hence non-inferiority (NI) trials are used which compare the experimental treatment with an active control. Such NI trials do not provide direct evidence on the superiority of the experimental treatment against placebo, and therefore historical information on placebo and active control treatments is used to estimate the placebo effect and define the NI margin. When the experimental treatment is a combination therapy, then often only limited or no information on the putative placebo is available, which requires more complex meta-analytic methods to estimate the placebo effect. The approach is illustrated by a case study in kidney transplantation, where a pivotal NI trial compared an active control treatment consisting of four immunosuppressant drugs, with an experimental treatment where two of these immunosuppressants were changed. As no historical information on the putative placebo treatment was available, a random-effect logistic meta-regression was used to estimate the placebo effect, and define the NI margin, based on data from 51 randomized controlled trials with a total of 17002 patients.

Issues in Determining Noninferiority Limits in Context with Retrospective Comparisons to Placebo

✦Steven Julious, University of Sheffield, Medical Statistics Group, ScHARR, Regent Court, 30 Regent Street, Sheffield, S14DA England, *s.a.julious@sheffield.ac.uk*

Key Words: non-inferiority, non-inferiority margin, fda, chmp

To assess the efficacy of a new investigative treatment a non-inferiority study is undertaken when it is no longer ethical to have a placebo control. Instead an active controlled trial is undertaken. The objective is to show that the new treatment is no worse than the active control A major issue in designing a non-inferiority study is the setting of the non-inferiority limit. The FDA discuss setting a limit so it would be possible to demonstrate superiority over placebo. This comparison would need to be done indirectly as placebo is not given concurrently. There is an issue with indirect comparisons if they are done retrospectively as the effect over placebo may not be as great today as when a placebo controlled trial was last undertaken The CHMP highlight the possibility of using a survey of experts to help quantify the margin. This could overcome some problems of retrospective comparisons - but can the actual empirical evidence be ignored? An approach that uses both the objective observed data and subjective opinion to set a non-inferiority margin would useful. This presentation explores setting a non-inferiority limit and discusses how simple Bayesian approaches may be useful

Methods for Determination of Noninferiority Margin

✦H.M. James Hung, Food and Drug Administration, 10903 New Hampshire Avenue, WO21, Room 4616, Silver Spring, MD 20993-0002, *hsienming.hung@fda.hhs.gov*

Key Words: non-inferiority margin, active control

The FDA non-inferiority draft guidance lays out a fundamental framework for design consideration in planning a non-inferiority trial or more broadly an active control trial. A key design specification in that framework is non-inferiority margin. The guidance contains a number of case examples showing margin determination under different scenarios. The statistical literature contains a number of methods for determining the margin. This presentation will shed some insights into the essential considerations behind these methods. A new method will be introduced for discussion.

567 Methods to address rare events, repeat events and regulatory requirements ■

Biopharmaceutical Section, Biometrics Section, ENAR, Section on Health Policy Statistics, Section for Statistical Programmers and Analysts, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Comparison Of Generalized Models For Analysis Of Standardized Incidence Rates Of Recurrent Events

✦John C Evans, Boston Scientific Corp, 01752 USA, *John.Evans@bsci.com*; Vivek Pradhan, Boston Scientific Corp; JIan Huang, Boston Scientific; Wen Ding, Boston Scientific

Key Words: Recurrent events, Standardized incidence rates, Patient years, Varying rate parameter

FDA has requested the standardized incidence rate of recurrent events per 100 patient-years for recent annual progress reports on mandated post-market studies. We consider two case studies, one involving a single device, and the other comparing two devices. Evans et al (2009) considered the distribution of the number of recurrent events per subject as either Poisson or Negative Binomial. They assumed that either events were randomly distributed over time or the rate parameter was constant over time. However, when the recurrence of events depends on the number and times of previous events, this can lead to a varying rate parameter. We will consider departures from these assumptions and compare the goodness of fit of counts to the Poisson and Negative Binomial distributions with other models, including the Generalized

Poisson, Double Poisson, Zero-Inflated Poisson, Zero-Inflated Negative Binomial, Poisson Inverse Gaussian, and Generalized Negative Binomial. We will compare results of analyses of standardized incidence rates based on these models.

Issues And Some Analysis Methods On Composite Endpoints

✦Nelson Lu, FDA/CDRH, 10903 New Hampshire Avenue, Silver Spring, MD 20993, *nelson.lu@fda.hhs.gov*; Rajesh Nair, FDA/CDRH; Chenguang Wang, FDA/CDRH; Yunling Xu, FDA/CDRH

Key Words: composite endpoint

A composite endpoint consists of multiple single endpoints that are combined in order to obtain a larger number of events during a clinical study. By setting up a composite endpoint, less sample size is required and thus the study is less costly. However, issues arise and are well recognized for such an endpoint. Oftentimes, the importance to the patient of the individual components may not be similar. Results may be dominated or driven by less important endpoint(s). The situation becomes even more challenging when there are competing risks between endpoints. Analyzing all individual components separately may partly address this issue. Yet it may not be able to provide convincing evidence due to insufficient power and maybe multiplicity issues. In this presentation, several methods to analyze a composite endpoint are proposed. Examples will be used to illustrate and assess these methods.

A Sensitivity Analysis Of Repeat Revascularization Causal Effect On Mortality And Morbidity In Drug Eluting Stent Clinical Trials

✦JIan Huang, Boston Scientific, 4100 Hamline Avenue North, South Town Square MS 9-315, St. Paul, MN 55112, *jhuangmd@yahoo.com*

Key Words: Causal effect, Sensitive analysis, Bias

In drug eluting stent clinical trials, Repeat Revascularization is often defined as an important safety outcome. However the revascularization procedure itself is cosidered a treatment. The usual approach for the estimation of treatment effect on mortality and morbidity may be biased, especially when many patients receive repeat revascularization due to disease progression after randomized for the original revascularization treatment. In this presentation, the different statistical methods on how to address the causal effect as well as real-life examples will be provided.

Assessment Of Safety Performance Goal Related To Rare-Event Meta-Analysis In Medical Device Studies

✦Hsini Terry Liao, Boston Scientific Corporation, 100 Boston Scientific Way, Marlborough, MA 01752-1234, *terry.liao@bsci.com*

Key Words: meta-analysis, rare event, sample size justification, performance goal, medical device

The study-level meta-analysis is sometimes involved in the sample size justification in terms of non-inferiority margin to performance goal in a single-arm study or to treatment effect in a two-arm clinical trial. In meta-analysis, there are challenges when at least one of the following quantities is small: sample sizes from individual studies, total number of studies and total number of events. Due to frequent innovations for specific indications in medical device world, low patient enrollment, lack of publications, and rare events have often been aware during study designing phase. Without loss of generality, assessment of the focus is on sensitivity of the performance goal in a single-arm study with those challenges being incorporated.

An Adaptive Strategy In A Single Arm Trial Against An Objective Performance Criterion Using An Exact Test Through The Binomial Distribution

Jin Wang, Abbott Vascular; ✦xiaolu su, abbott vascular, 3200 lakeside drive, santa clara, CA 95054, *xiaolu.su@av.abbott.com*

Key Words: adaptive design, conditional error function, binomial distribution, single-arm trial

For medical product development within the same generation, usually a small incremental change is made. Therefore, single-arm trial designs are commonly implemented to test the performance of the new product against an Objective Performance Criterion (OPC), which is a fixed value developed by historical data or literature review. When the primary endpoint is an event rate, to be conservative, regulatory agencies may require the use of an exact test through binomial distribution when the sample size is modest. Utilizing the idea by Proschan and Hunsberger, this article proves that it can be a free gift to add an adaptive component to a regular fixed sample size design so that when the result in the first stage is marginal, the adaptive feature can be activated without any penalty. A hypothetical example is used to illustrate the application of this method.

568 Large scale genetic studies of complex traits ■●

ENAR, International Indian Statistical Association, Section on Statistics in Epidemiology, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

The Linkage Disequilibrium Lasso For Snp Selection In Genetic Association Studies

✦Samuel G. Younkin, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106 USA, *samuel.younkin@gmail.com*; J. Sunil Rao, University of Miami; Robert C Elston, Case Western Reserve University; Joseph Nadeau, Institute for Systems Biology

Key Words: LASSO, GWAS, Alzheimer's, Linkage Disequilibrium, R, SNP

A rapid increase in the number of available human genetic variants has led to an agnostic approach to disease-gene mapping which has resulted in the popularization of the genome-wide association study. This method has identified only a small portion of the expected disease susceptibility sites, and it is our belief that many of the remaining SNPs contribute such a small amount to overall heritability that studies cannot overcome problems due to multiple testing. Here we develop a statistical method designed for a genetic association signal that is more representative of the remaining disease susceptibility SNPs. The genetic association signal that we seek is no longer a peak, for the increase in SNP density, coupled with low effect sizes give rise to a plateau-like signal with gaps. We address this by formulating our method as a penalized least squares regression estimator based on the linkage disequilibrium present between SNPs. The method known as the LD LASSO is an adaptation of the fused LASSO used for subset selection when the signal is sparse and block-like. We implement this method in the R package ldlasso, and present results using data from an Alzheimer's GWAS.

A Discussion Of Current Applications And Statistical Issues In Genomic Prediction Of Complex Traits And Diseases

✦Christine Woods Duarte, University of Alabama at Birmingham, RPHB 327, 1530 3rd Ave. S., Birmingham, AL 35294-0022, *cduarte@uab.edu*; Gustavo de los Campos, University of Alabama at Birmingham

Key Words: complex traits, GWAS, prediction, QTL, WGP

High dimensional genomic techniques have been used widely in recent years for association studies and for prediction of complex traits and diseases. Specific applications include genome-wide association studies (GWAS) for complex human diseases and genomic selection (GS) in animal and plant breeding. Prediction of complex traits and phenotypes using genetic data is an important goal, and while significant advances have been made, room for improvement remains before goals such as personalized medicine can be achieved. At the crux of the issue are statistical methods for creating prediction models from high-dimensional data. Whole-genome prediction models (WGP) have been highly successful in animal breeding, and some initial work has been done in the context of human disease. Techniques using variable selection from GWAS studies have been explored as well. We will discuss genetic and statistical factors affecting prediction accuracy (sample size, extent of linkage disequilibrium, genetic architecture and trait heritability), address ways of confronting complexity (gene x gene and gene x environmental interaction) and discuss advantages/disadvantages of model selection versus WGP.

Machine Learning Techniques For Predictive Modeling In Pharmacogenetics: Modeling Warfarin Dose Response In African Americans

✦Erdal Cosgun, University of Alabama at Birmingham, Department of Biostatistics 1665 University Boul., Ryals Public Health Building 414, Birmingham, AL 35294, *erdalcosgunn@gmail.com*; Nita A. Limdi, University of Alabama at Birmingham; Christine Woods Duarte, University of Alabama at Birmingham

Key Words: Machine Learning, Warfarin, Data Mining, Random Forest Regression, Support Vector Regression, Boosted Regression Tree

With most complex traits and diseases having expected genetic contributions of many hundreds or even thousands of genetic factors,and with genotyping arrays consisting of hundreds of thousands SNPs,powerful high dimensional statistical techniques are needed to comprehensively model the genetic variance.We have previously applied 3 machine learning (ML) approaches:Random Forest Reg.,Boosted Reg. Tree and Support Vector Reg. to the problem of prediction of warfarin maintenance dose in a sample of African Americans. We have developed a multi-step approach for SNP selection,model building, and model assessment in a cross-validation framework,and our results indicate that our modeling approach gives much higher accuracy than previous models for warfarin dose prediction,with an average R2 between predicted and actual square root of warfarin dose as measured in the test samples ranging from 52.4% to 68.2% depending on the method used.We are currently investigating larger prediction models with up to a thousand pre-selected SNPs,improved methods for marker pre-selection.In summary,ML approaches for high-dimensional pharmacogenetic prediction hold great promise and warrant further research.

Predicting Complex Human Traits Using Whole-Genome Markers: Proof Of Principle With Human Height And Human Longevity

✦Gustavo de los Campos, University of Alabama at Birmingham, RPHB 327, 1530 3rd Ave. S., Birmingham, AL 35294-0022, *gdeloscampos@ms.soph.uab.edu*; Yann Klimentidis, University of Alabama at Birmingham; Robert Makowsky, University of Alabama at Birmingham; Ana Ines Vazquez, University of Alabama at Birmingham; Nicholas M Pajewski, - Department of Biostatistical Sciences, Wake Forest University School of Medicine ; Christine Woods Duarte, University of Alabama at Birmingham; David B. Allison, University of Alabama at Birmingham

Key Words: Prediction, SNPs, Complex Traits

Most diseases have a genetic component. However, despite great progress in genotyping technologies, our ability to predict genetic risk remains limited. A perhaps overlooked problem resides in limitations of the statistical methods used in genome-wide association studies. These methods, based on single marker associations, are most powerful for traits affected by a few genes, but are not well suited to complex traits. We applied whole-genome prediction (WGP), a predictive approach which uses dense single nucleotide polymorphism (SNPs) data, to predict human height (H) and years of life (YL). Data was from the Framingham Heart Study (N=5,117). Models with different marker density (from 2.5K to 400K SNPs) were fitted using different WGP methods. Using 400K SNPs we achieved a cross-validation-R2 (CV-R2) of 0.25 for (age and sex-adjusted) H. Previous studies with H have failed to achieve an R2>0.10 for H. For YL we achieved a CV-R2 of 0.21. The contribution of SNPs to CV-R2 in predicting YL was similar to that of sex, smoking and body-mass index combined. We conclude that WGP can offer opportunities to advance our ability to predict complex human traits.

A Two-Step Modeling Strategy For Testing And Estimating Genetic Susceptibility To The Ill-Effects Of Adiposity: Illustration In An Outbred F2 Mice Population

✦Guo-Bo Chen, Section on Statistical Genetics, Dept. of Biostatistics, Univ. of Alabama at Birmingham, RPHB 327, 1530 3RD AVE S, Birmingham, AL 35294-0022 United States, *bchen@ms.soph.uab.edu*; Katherine H Ingram, Section on Statistical Genetics, Dept. of Biostatistics, Univ. of Alabama at Birmingham; Gustavo de los Campos, University of Alabama at Birmingham; Nengjun Yi, Section on Statistical Genetics, Dept. of Biostatistics, Univ. of Alabama at Birmingham; Xiang-Yang Lou, Section on Statistical Genetics, Dept. of Biostatistics, Univ. of Alabama at Birmingham; Daniel Pomp, Dept. of Genetics, University of North Carolina; David B. Allison, University of Alabama at Birmingham

Key Words: Prediction, Bayesian Lasso, insulin resistance, Linear regression, Interaction, Adiposity

We explored the possibility of using genome-wide data to achieve the types of predictions that would be useful in personalized medicine. We examined an outbred F2 mouse population derived from an intercross between the line M16 and the line ICR. Specifically, we asked “Can we identify mice for whom their degree of fatness will strongly influence their degree if insulin resistance and those for whom their degree of fatness will be unrelated to their degree if insulin resistance”. To illustrate how the variation in our phenotype (Y) in question, insulin sensitivity, is determined by body composition (P) and genes (G) as well as their interaction, a two-step statistical model was proposed. In step one, the genetic values associated with a joint measure of P and Y are inferred by regressing phenotypes on all markers concurrently, with their effects estimated using the Bayesian Lasso. In step two, Y is regressed on the predicted genetic values (from step one) and percent fat and their interaction (B). The p-value for the interaction effect is most significant ($p< 0.001$). These results underscore the importance of both body composition and genetic influence on insulin resistance.

569 Bayesian Nonparametric and Semiparametric Methods 2

Section on Bayesian Statistical Science, Section on Nonparametric Statistics

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

A Bayesian Hierarchical Model For Estimating And Partitioning Bernstein Polynomial Density Functions

✦Charlotte C. Gard, Department of Biostatistics, University of Washington, Seattle, WA , *gardc@u.washington.edu*; Elizabeth R. Brown, Department of Biostatistics, University of Washington

Key Words: Bayesian hierarchical models, Bernstein polynomials, mixture models, partitioning, density estimation, birth-death MCMC

We present a Bayesian hierarchical model for simultaneously estimating and partitioning probability density functions. Individual density functions are modeled using Bernstein densities, which are mixtures of beta densities whose parameters depend only on the number of mixture components. We place a prior on the number of mixture components and write the mixture weights as increments of a distribution function G. We place a Dirichlet process prior on G and allow the parameters of the Dirichlet process, the baseline distribution and the precision parameter, to be random. We use a mixture of (a product of) beta densities to partition subjects into groups, with subjects in the same group sharing information via a common baseline distribution. Inference is carried out using Markov chain Monte Carlo. We offer a computing algorithm based on the constructive definition of the Dirichlet process. We consider the case where the number of groups is fixed and the case where the number of groups is unknown, using a birth-death algorithm to make inference regarding the number of groups. We demonstrate the model using radiologist-specific distributions of percent mammographic density.

Simple Outlier Detection with Dirichlet Process Mixtures

✦Matthew Shotwell, Vanderbilt University School of Medicine, Department of Biostatistics, 1161 21st Ave South, S2323 MCN, Nashville, TN 37232, *matt.shotwell@vanderbilt.edu*; Elizabeth H Slate, Medical University of South Carolina

Key Words: outlier, product partition model, Dirichlet process mixture, Bayes factor, MAP estimate, optimization

Title: Simple Outlier Detection with Dirichlet Process Mixtures We introduce a Bayesian inference mechanism for outlier detection using the product partition model equivalent of a Dirichlet process mixture. Outliers are detected by forming a maximum a posteriori (MAP) estimate of the data partition. Observations that comprise small or singleton clusters in the estimated partition are considered outliers. We present a novel interpretation of the Dirichlet process precision parameter and demonstrate its utility in outlier detection problems. The precision parameter is used to form an outlier detection criterion based on the Bayes factor for a partition with outliers versus a class of partitions with fewer or no outliers. We also introduce a computational method for MAP estimation that is free of posterior sampling and guaranteed to find a MAP estimator in finite time. Our computational method is compared with several established methods in a real data example. Implementation of the methodology for product partitions of linear models is available in the R package profdpm.

A Bayesian nonparametric modeling framework for developmental toxicity studies

✦Kassandra Fronczyk, University of California Santa Cruz, , *kassief@ams.ucsc.edu*; Athanasios Kottas, University of California, Santa Cruz

Key Words: Dependent Dirichlet process, Developmental toxicology data, Dirichlet process mixture models, Gaussian process, Markov chain Monte Carlo, Risk assessment.

We develop a Bayesian nonparametric mixture modeling framework for replicated count responses in dose-response settings. We explore this methodology for modeling and risk assessment in developmental toxicity studies, where the primary objective is to determine the relationship between the level of exposure to a toxic chemical and the probability of a physiological or biochemical response, or death. Data from these experiments typically involve features that can not be captured by standard parametric approaches. To provide flexibility in the functional form of both the response distribution and the probability of positive response, the proposed mixture model is built from a dependent Dirichlet process prior, with the dependence of the mixing distributions governed by the dose level. The methodology is tested with a simulation study, which involves also comparison with semiparametric Bayesian approaches to highlight the practical utility of the dependent Dirichlet process nonparametric mixture model. Further illustration is provided through the analysis of data from two developmental toxicity studies.

Single Factor Transformation Priors For Density Regression

✦Suprateek Kundu, Biostatistcs, UNC Chapel Hill, 3101 McGavran-Greenberg Hall, CB #7420, CHAPEL HILL, NC 27599, *skundu@email.unc.edu*; David Dunson, Duke University

Key Words: Nonparametric Bayes, Kernel estimation, Density regression, Gaussian process, Latent variable model, Dirichlet process

Although mixture modeling has formed the backbone of the literature on Bayesian density estimation incorporating covariates, the use of mixtures leads to some well known disadvantages. Avoiding mixtures, we propose a flexible class of priors based on a random transformation of a uniform latent variable. These priors are related to Gaussian process latent variable models proposed in the machine learning literature. For density regression, we model the response and predictor means as distinct unknown transformation functions dependent on the same underlying latent variable, thus inducing dependence through a single factor. The induced prior is shown to have desirable properties including large support and posterior consistency. We demonstrate advantages over Dirichlet process mixture models in a variety of simulations, and apply the approach to an epidemiology application.

Posterior Consistency In Conditional Distribution Estimation

✦Debdeep Pati, Department of Statistical Science, Duke University, Durham, NC 27708, *dp55@stat.duke.edu*; David Dunson, Duke University

Key Words: Asymptotics, Bayesian nonparametrics, Density regression, Large support, Probit stick-breaking process, Dependent Dirichlet process

A wide variety of priors have been proposed for nonparametric Bayesian estimation of conditional distributions, and there is a clear need for theorems providing conditions on the prior for large support and posterior consistency. Estimation of an uncountable collection of conditional distributions across different regions of the predictor space is a challenging problem, which differs in some important ways from density and mean regression estimation problems. Focusing on a broad class of priors formulated as predictor-dependent mixtures of Gaussian kernels, we provide sufficient conditions under which weak and strong posterior consistency hold. This theory is illustrated by showing that the conditions are satisfied for a class of generalized stick-breaking process mixtures in which the stick-breaking lengths are constructed through mapping continuous stochastic processes to the unit interval using a monotone differentiable link function. Probit stick-breaking processes provide a computationally convenient special case. We also provide a set of sufficient conditions to ensure strong and weak posterior consistency using fixed-π dependent Dirichlet process mixture of Gaussians.

570 Statistical Modeling and Analysis in Financial Industries ■

Section for Statistical Programmers and Analysts, International Chinese Statistical Association, Section on Statistical Computing, Section on Statistical Graphics, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Multiple Failure Time Prediction By Discrete Time Survival Model

Li Yang, China; ✦Limin Yang, Renmin University, , *yanglimin@ruc.edu.cn*

Key Words: Predictive Modeling, Discrete-Time, Survival Model, Multiple Failure Time, Prediction

Predictive modeling becomes more and more popular and important. We are not only need to describe the data, but also need to predict the future. Most of the current studies are focusing on the models for descriptive purpose. The big difference between descriptive model and predictive model is that descriptive model can use any variables as long as they make sense, but predictive model can only use those variables available at predictive time. In order for predictive model to be feasible, data structure and variable definition have to be selected and defined clearly, especially for those time dependent variables. Discrete-time survival model is often used to model the failure time or duration such as customer attrition time, loan default time, new product buying time, etc. In this paper, we develop a new model based on discrete time survival model. We will show how to extend the descriptive model for predictive purpose. That is to predict multiple time failure by using discrete-time survival model, instead of one time. And we also create a conditional logit for multiple failure time prediction purpose. An example is also shown.

Why Adam (Analysis Data Model)?

✦Kevin Lee, Cytel, 640 Lee Road, Suite 201, Chesterbrook, PA 19087, *kevin.lee@cytel.com*

Key Words: ADaM, CDISC, Time to Event, BDS, Statisticians, SAS

The pharmaceutical industry is embracing CDISC as the standard of clinical trial process and ADaM is the part of the process that statisticians and clinical trial programmers is developing and using to analyze the clinical trial data. But, one can ask why ADaM? Why do we need to create ADaM? What are the benefits to use ADaM Data Structure? How can statisticians use them? The paper is intended for statisticians who are interested in ADaM and also have above questions. The paper will discuss the data structure of ADaM ADSL and ADaM BDS (Basic Data Structures). It will also discuss the role of ADaM in CDISC and its benefits in clinical trial. It will also discuss the roles of Clinical Trial SAS programmers and statisticians in ADaM (Usually, it depends on the company). The paper will provide the example of ADaM Data Set (Time to Event) and SAS codes that generate the final numbers in the analysis. The example will display how ADaM Data set is built and how it can be used in the analysis.

Estimation Of Stock Market Crashes On The Basis Of Multivariate Skew-Student Density

✦Lei Wu, School of Economy,Nankai University, School of Economy, Nankai University, Tianjin, 300071 China, *shalinnagr@gmail.com*; Qingbin Meng, School of Business, Renmin Universty ; Julio Velazquez, Applied Mathematics, Delft University of Technology

Key Words: Multivariate skew-Student density, TVC-GARCH model, Stock market crashes, Crash probability forecast

By combining the multivariate skew-Student density with a time-varying correlation GARCH (TVC-GARCH) model, this paper investigates the spread of crashes in the regional stock markets in both bilateral and global environments. This empirical framework provides a more accurate specification for empirical return distributions and helps us to detect the dynamic properties of crashes between regions. By computing the conditional 1-day crash probabilities and evaluating the forecast performance of the TVC-GARCH model with different densities, it is concluded that the multivariate skew-Student density can explain substantial dependence in the occurrence of crashes regions and predict global crashes with better predictive accuracy only depending on the past information set. This empirical framework also gives insight into the further research about contagion and can be used to improve the early-warning systems.

Statistics Methods In The Risk Management Of Bank

Chang Song, Renmin University of China; ✦Jie Zhang, Renmin University of China, Beijing, 100872 China, *jacqueline35@sina.com*

Key Words: risk management, bank governance, value at risk, China, financial crisis

This paper examines the impact of bank governance on risk management by using the statistics method. Some studies address the question of how bank governance practices affect risk management. However, an application of the statistic method for empirical research to this issue has received less attention, especially in emerging markets. Our dataset includes a sample for all listed commercial banks in China from year 2007 to 2009. The results suggest that bank governance is relevant to risk management. Specifically, the shareholding of the largest shareholder affects bank risk management but its effect is non-linear, but little evidence in executive compensation and other factors of corporate governance. The banks may actually benefit from the good bank governance, which can improve the risk management, especially during financial crisis.

The Statistical Properties And Modeling Of The Spread Decomposition On Emerging Market

Limin Yang, Renmin University; ✦YIhu Qin, Renmin University of China, , *qinyihu2000@yahoo.com.cn*; Yan Li, Renmin University of China

Key Words: Spread Decomposition, Emerging Market, Statistical Properties

Spread, as a measure of stock market implicit trading costs, are closely related with stock market liquidity, volatility, efficiency, transparency and other features. It can reflect the level of stock market efficiency. The smaller the spread is, the higher the efficiency of security market is. To better understand the spread formation, we generally need to model spread decomposition. China, as an emerging market, has greater different spread composition than mature markets. In this paper, we study the spread decomposition on emerging market based on the event that how the three adjustments on the usage of stamp duty impact the spread. First, we studied the statistical properties of spread changing between before the stamp duty adjustments and after the stamp duty adjustments. Second we used the LSB model to decompose the spreads and compare ratios changed of every component of the spread between before the stamp duty adjustments and after the stamp duty adjustments. Finally, we used EROP model to validate the changes of the trader proportions.

An Adam-Based Analysis & Reporting Knowledge Library

✦Thomas W Dobbins, Merck & Co., Inc., 351 N. Sumneytown Pike, UG1D-80, North Wales, PA 19454-2505, *thomas_dobbins@merck.com*; John Troxell, Merck & Co., Inc.; Ram Radhakrishnan, Merck & Co., Inc.

Key Words: CDISC, ADaM, Analysis & Reporting, Knowledge Library

Any organization that routinely performs analysis and reporting (A&R) of data can benefit from a holistic approach to building and maintaining analytical knowledge to support this A&R. In this paper, the concept of a Knowledge Library (KL) is introduced that encompasses connected, modular standards for analysis data and metadata structures, statistical methods, and analysis displays (tables and figures). This KL can be the basis for operational efficiency, and through SAS dataset standards compliant with an industry standard such as the Analysis Data Model (ADaM) of the Clinical Data Interchange Standards Consortium (CDISC), this KL also provides key building blocks to enable the ongoing accumulation and preservation of knowledge in an accessible way for both immediate and future use. The authors present and discuss the concept and design considerations for a KL in a pharmaceutical research setting where a KL can efficiently address both short term needs, such as scientific discovery and submission to regulatory agencies, and longer term use cases, such as queries, data mining, integrated analysis, and meta-analysis.

A Scientific Obligation: Supporting The Research Process With Complete, Documented Data

✦Steven Kirby, Celerion, 621 Rose St., Lincoln, NE 68502, *steven.kirby@celerion.com*; Aleksandra Stein, Celerion

Key Words: Data Standards, Metadata, CDISC, SDTM, ADaM

Conclusions not supported by data are just opinions. Scientists in many disciplines are working to generate standards for how scientific data should be delivered to end users. As scientists who support subject matter experts across all scientific disciplines, Statisticians are uniquely qualified to help the scientific community better define what makes data useful and accessible. With data more easily collected, shared and analyzed than ever, it is important for statisticians to help lead the discussion on how best to form general requirements for scientific data. The increasingly formalized data requirements associated with the US drug development process (based on FDA guidance and data standards from the CDISC consortium) will be reviewed to highlight basic characteristics of data that solidly support scientific conclusions. How standards for the delivery of scientific data and associated documentation have been practically implemented across disciplines will be investigated, and the strengths and weaknesses of existing approaches will be evaluated. We conclude that education and outreach on general standards for useful data is needed; possible approaches are discussed.

571 Modern Methods for Inferring About Treatment Effects

Biometrics Section, Biopharmaceutical Section, ENAR

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Stratified Randomization With Gender-Balancing In The Sure-Pd Trial

✦Eric A Macklin, Massachusetts General Hospital, 50 Staniford St, Suite 560, Boston, MA 02114, *emacklin@partners.org*; Michael A Schwarzschild, Massachusetts General Hospital; Rebecca A. Betensky, Harvard School of Public Health, Harvard University

Key Words: Randomization, institutional balancing, stratification, clinical trial design, Parkinson's disease

Randomization designs aim to assign treatments in appropriate proportions while minimizing confounding. In multicenter trials, clinical site can be an important confounder, and randomization is typically stratified by site. When expected enrollment at each site is low, further stratification may not be practical. Study-wide balancing permits balanced treatment allocation marginally for additional potential confounders. We report on four alternative randomization designs for a three-arm multicenter trial in Parkinson's disease: stratification by site, stratification by site and gender, stratification by site with gender balancing, and stratification by gender with site balancing. We compare these designs through simulation with respect to: overall treatment balance, balance by site, balance by gender, drug supply, and power for treatment main effect, interactions, and gender-specific contrasts. We find that stratification by site with gender-balancing is not superior to stratification by both site and gender even with limited enrollment per site. We discuss criteria for choosing among alternative randomization designs for multicenter trials with several potential confounders.

Impact Of Sample Size On Prognostic Imbalance And Covariate Adjustment In Randomized Controlled Trials (Rcts)

✦Rong Chu, McMaster University, Health Sciences Centre, Room 2C7, 1200 Main Street West, Hamilton, ON L8N 3Z5 Canada, *chur@mcmaster.ca*; Stephen Walter, McMaster University; Lehana Thabane, McMaster University; Philip James Devereaux, McMaster University; Gordon Guyatt, McMaster University

Key Words: Randomized Controlled Trials, Prognostic Imbalance, Covariate Adjustment, Simulation Study, Sample size

Randomization does not guarantee balance in all baseline covariates particularly in trials with small sample sizes. Study objectives include: 1) To assess the probability of imbalance in a rare yet strong prognostic factor (PF) between two arms; 2) To investigate the impact of prognostic imbalance on statistical modeling of treatment effect; 3) To examine the effect of sample size in relation to the first two objectives. We simulated RCTs with a binary outcome by varying the risk of outcome in the control group, effect of treatment, effect of PF, prevalence of PF, and sample size. Logistic regression models with and without adjustment for PF were compared in terms of bias, standard error, coverage of confidence interval and statistical power. The probability of imbalance increase as sample size decreases. For a PF with prevalence of 0.5, the probability of an imbalance = 5% can reach 0.4 with 125 per arm. Ignoring a strong PF (relative risk = 5) leads to bias in assessing moderate treatment effect; the bias is essentially independent of sample size. Adjusting for PF reduces precision but increases statistical power. Results conditional on 5% imbalance will be presented.

An Information Criterion For Sensitivity Analyses Of The Treatment Ignorability Assumption

✦CHEN-PIN WANG, UTHSCSA, TX 78229, *WANGC3@UTHSCSA.EDU*

Key Words: causal modeling, treatment ignorability, information criterion, Kullback-Leibler Distance

The treatment ignorability assumption (Rubin 1978) is critical in causal modeling. This paper proposes a Kullback-Leibler Distance based information criterion (KLD-IC) suitable for conducting sensitivity analyses for verifying the treatment ignorability assumption in the causal modeling framework. Here we focus on how the inference of certain statistic T_{n} is affected by the violation of the treatment ignorability assumption. Stemmed from Goutis and Robert (1998), our proposed KLD-IC is the posterior mean of the KLD between the (predictive) distribution of a statistic T_{n}, under two likelihoods r and f, where r is the likelihood when the ignorability assumption is met and f is the the likelihood when the ignorability assumption is violated. We examine the asymptotic properties of this KLD-IC under certain regularity conditions. We also show the impact of this asymptotic property for examples where only part of the regularity conditions are satisfied. We applied this KLD-IC to a study of pain among those with opioid dependence.

Group Sequential Methods Based on Ranked Set Samples

✦ESAM ABDUL QADER AL-SAWI, King Fahd University for Petroleum and Minerals, Saudia Arabia, Dhahran, Dhahran, International +966 KSA, *alsawies@kfupm.edu.sa*; Abdul K. Hussein, University of Windsor; H.A. Muttlak, King Fahd University of Petroleum and Minerals

Key Words: Auxiliary information, Group sequential methods, Ranked set sampling, Two-sample hypotheses

Ranked set sampling schemes were originally proposed to increase efficiency in estimation. On the other hand, group sequential methods provide substantial savings in sample and enable us to make decisions as early as possible. In this manuscript, we intend to combine the benefits of the two methodologies. We propose group sequential tests for one and two population means under ranked set sampling. We compare the power, average sample sizes and type I errors of the proposed tests

to those of group sequential tests based on simple random sampling schemes. We illustrate the utility of the method by using data from a leukemia clinical trial.

Estimating Treatment Effects From A Randomized Trial In The Presence Of Post-Study Treatment

✦Min Zhang, University of Michigan, M4126 SPH II, 1415 Washington Heights, Ann Arbor, MI 48109-2029, *mzhangst@umich.edu*; Yanping Wang, Eli Lilly and Company

Key Words: Causal inference, Inverse probability weighting, Covariate adjustment, Cancer, Marginal structural model, Clinical trial

In randomized clinical trials involving survival time, a challenge that arises frequently, for example in cancer studies, is that during follow up subjects may initiate post-study treatment (PST), which is not part of the study design and could potentially confound effect of the study drug. Marginal structural Cox's model and methods based on inverse probability weighting have been proposed to account for PST in the presence of time-dependent confounders. Inverse probability weighting methods tend to yield estimators that are of big variance and not stable. These methods were originally proposed for observational data and do not target the question of interest. In this paper, we adopt the marginal structural Cox's model and propose an augmented inverse probability weighting method. The proposed method improves the efficiency of the usual methods by exploiting the fact that the study treatment is independent of baseline covariates, guaranteed by randomization. The finite-sample performance of the proposed method is demonstrated via simulation studies and by application to data from a cancer clinical trial.

Competing-Risk Survival Analysis Of Bariatric Surgery Data Using Propensity-Score Methods

✦Diqiong Xie, University of Iowa, 611 Hawkeye Drive, Iowa City, IA 52246 US, *diqiong-xie@uiowa.edu*; Michael P Jones, University of Iowa; Mason E Edward , University of Iowa Hospitals and Clinics

Key Words: propensity score, matching, survival analysis, competing risks, bariatric surgery

Bariatric surgery has been recognized as an effective therapy for morbid obesity. There are currently two major categories of surgical procedures, depending on whether gastric bypass is included or not. 'Simple' operations are purely restrictive, with no bypass of digestive tract, while 'Complex' operations bypass the duodenum and/or jejunum, in addition to decreasing the size of the stomach. There have been several observational studies investigating the difference in outcome between these two types of procedures, focusing on the efficiency of weight loss, post-operative complication and overall mortality. However, the question remains unanswered about how selection bias was adjusted in the above mentioned observational studies, in which the treatments, i.e. operation types, were not randomly assigned, but were determined by the surgeon and patient. In this study, we show that the choice of treatment depends heavily on patient characteristics. Therefore, we adjust for selection bias by using propensity scores, and analyze the effect of operation types on overall mortality and cause-specific mortality using pair-matched patients.

Weighted Cumulative Treatment Estimation For Sequentially Randomized Clinical Trials In The Presence Of Non-Proportional Hazards

✦Xinyu Tang, University of Arkansas for Medical Sciences, , *xtang@uams.edu*; Adbus S. Wahed, University of Pittsburgh

Key Words: non-proportional hazards, cumulative baseline hazard, stratified proportional hazard model, treatment regime

Proportional hazards model is widely used in survival analysis to adjust for baseline covariates. Although Lokhnygina and Helterbrand (2007) introduced Cox regression methods for two-stage randomization designs, their method can only be applied to test the equality of two treatment regimes that share the same maintenance therapy. Moreover, their method does not allow auxiliary variables to be included in the model. Besides, in many medical studies, the treatment effect is not constant over time. Wei and Schaubel (2008) proposed an estimator of cumulative treatment effects based on treatment-specific cumulative baseline hazards using a stratified proportional hazards model. Their model assumed proportionality within treatment group and non-proportionality across treatment groups. We propose a similar approach to estimate the cumulative treatment effect for treatment regimes from sequentially randomized clinical trials. Comparisons among treatment regimes are performed by testing the ratio of the estimated cumulative hazards. A simulation study was conducted to evaluate the performance of the estimators and proposed tests.

572 Bayesian methods for rare events, sparse signals, and change points

Section on Bayesian Statistical Science, ENAR, Section on Statistics and the Environment, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

On Threshold Estimation In Threshold Regression Models

✦Friederike Greb, University of Goettingen, Wilhelm-Weber-Strasse 2, Goettingen, International 37073 Germany, *fgreb@uni-goettingen.de*; Tatyana Krivobokova, Georg-August-Universitaet Goettingen; Axel Munk, University of Goettingen; Stephan von Cramon-Taubadel, University of Goettingen

Key Words: threshold regression model, threshold estimation, Bayesian estimator, empirical Bayes, nuisance parameters

Estimation of threshold parameters in various threshold regression models is typically performed by maximizing the corresponding profile likelihood. However, in certain situations such estimator performs poorly. For example, the threshold estimator can be seriously biased if the true threshold divides the data into sets of very unequal sizes and many nuisance parameters are present. A natural solution would be to employ a Bayesian estimator (BE). BEs have been considered in the literature. In particular, in an i.i.d. setting asymptotic optimality results, which are independent of priors, have been proven for BEs. Yet we find that the choice of priors can crucially affect estimation results in small samples and non-informative priors can distort estimates. In contrast, with an adequate parametrization and selection of priors we are able to regularize nuisance parameter estimates. Using an empirical Bayes method, we obtain the penalty coefficients in a data-driven manner. Simulation results show that this produces an estimator which performs well even in situations where commonly used estimators fail. We illustrate the relevance of our approach with several real-data examples.

Change-Point Analysis Using Shape Restricted Regression Splines In A Bayesian Framework.

✦Amber Hackstadt, Colorado State University, 720 City Park Ave, Apt C320, Fort Collins, CO 80521, *ahacksta@lamar.colostate.edu*; Mary C Meyer, Colorado State University; Jennifer Hoeting, Colorado State University

Key Words: change-point, shape-restricted regression, regression splines, free knot splines

We extend a Bayesian shape-restricted regression spline model to allow for one or more change-points. The Bayesian shape-restricted regression spline model uses fixed or free knot regression splines along with vague priors to estimate functions while providing posterior distributions that facilitate inference. The model allows for the selection of the location of the change-points while retaining a-priori information about shape. The Bayesian framework facilitates inference and the regression spline framework provides flexibility in function estimation.

A Bayesian Approach For Signal Detection In Noisy Images

Khalil Shafie, Shahid Beheshti University; ✦Mohammad Reza Farid Rohani, Shahid Beheshti University, 1983963113 Iran, *mf-rohani@sbu.ac.ir*

Key Words: Bayes factor, Gaussian random fields, signal detection, Hilbert valued Normal

Rohani et. al (2006) considered the problem of searching for activation in brain images obtained from functional magnetic resonance imaging (fMRI) and the corresponding functional signal detection problem. They developed a Bayesian procedure to detect signals existing within noisy images when the image is modelled as a scale space random field. The purpose of this paper is to extend the scale space result to a more general setting. In a genereal abstract setting, using Radon-Nikodym derivative, an extended definition of Bayes factor for testing the point null hypothesis is presented. Using this extended definition, a Bayesian testing procedure for signal detection in noisy images when both signal and noise considered as an element of an infinite dimensional Hilbert space is introduced. The method is applied to the problem of searching for activation in brain images obtained by functional magnetic resonance imaging (fMRI).

Incorporating Model Uncertainty In Detecting Rare Variants: The Bayesian Risk Index

✦Melanie Quintana, University of Southern California, Los Angeles, CA 90089 US, *wilsonme@usc.edu*; Duncan Thomas, University of Southern California; Jonine Berstein, Memorial Sloan-Kettering Cancer Center ; David V. Conti, University of Southern California

Key Words: rare variants, risk index, genetic association studies, Bayesian model uncertainty, multiplicity correction, sequence analysis

We are interested in investigating the involvement of multiple rare variants within a given region by conducting analyses of individual regions with two goals: (1) to first determine if regional rare variation in aggregate is associated with risk; and (2) conditional upon the region being associated we wish to identify specific genetic variants within the region that are driving the association. In particular, we seek a formal integrated analysis that achieves both of our goals. Our framework aims at constructing a risk index based on multiple rare variants within a region and then detecting associated regions based on the indices. Our analytical strategy is novel in that we use a Bayesian approach to incorporate model uncertainty in the selection of which variants to include in the index as well as whether the variant is involved as a protective or risk contributing factor. Using a set of study-based simulations, we show that our methodology has added power over the more commonly used methods. In addition, we apply the approach to sequence data within BRCA1 for samples from the WECARE Study of second primary breast cancers.

Performance Of Bayesian Ranking Methods For Identifying The Extreme Parameter

✦Yi-Ting Chang, Johns Hopkins University, , *yiting0823@gmail.com*

Key Words: Bayesian Ranking

The Bayesian approach provides a unified framework for estimating unit specific parameters in multilevel models. Several optimal estimators of ranks under different loss functions have been proposed, and performance evaluations have been conducted as well. However, in some cases, our focus is only on identifying the unit with the largest underlying value rather than the whole population. Thus, we compare the performances of the different rank estimators in the literature by simulations in terms of identifying the top-ranked (spiked) unit. We also apply these ranking methods to simulated genetic outcomes from a real data drawn from the International HapMap Project.

Bayesian Hierarchical Shrinkage Prior For Sparse Signals

✦Yingbo Li, Duke University, Department of Statistical Science,, Duke University, Durham, NC 27708 United States, *yl118@duke.edu*; Merlise Clyde, ISDS, Duke University

Key Words: Bayesian model averaging, variable selection, shrinkage, sparsity

In this paper, we propose a new fully Bayes approach for estimating sparse signals from background Gaussian white noise. We construct a hierarchical shrinkage prior as mixtures of Cauchy densities, which in the limit leads to a Levy random field prior. The prior can also be considered as a mixture of a point mass at zero and a heavy-tailed density which permits it to adapt to different sparsity structures. By studying the tail property of its induced marginal likelihood, we prove it has bounded influence. This hierarchical shrinkage prior shrinks small values directly towards zero while keeping large signals almost unshrunk. The infinite divisibility of our prior leads to coherent prior specifications as the number of predictors increases. Based on simulation studies, under certain circumstances, our prior can achieve higher accuracy in terms of sum of squared error, comparing with some existing Bayesian model selection approaches such as Johnstone and Silverman's Empirical Bayes estimates.

573 Statistical Applications of Longitudinal Methods

Biometrics Section, Biopharmaceutical Section, ENAR, International Indian Statistical Association, Section on Statistics in Epidemiology

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Applying Generalized Linear Mixed Models To Word Counts To Analyze The Literary Style Of Pre-1920 Detective Fiction

✦Roger Bilisoly, Dept. Mathematical Sciences, Central Connecticut State University, 1615 Stanley St. P.O. Box 4010, New Britain, CT 06050-4010, *bilisolyr@ccsu.edu*; Krishna K Saha, Central CT State University

Key Words: GLMM, Stylometry, Categorical Data, Text Mining

We investigate differences in literary style in pre-1920 detective short stories by A. C. Doyle, G. K. Chesterton, and E. W. Hornung, where generalized linear mixed models (GLMMs) are applied to word counts. To simplify the analysis, only a few restricted sets of words are considered, which form word classes that share both (1) a related meaning and (2) common grammatical features. For instance, one class is the "break verbs," which includes "break," "crack," and "crash," and these use similar verb alternations (types of grammatical structures) in sentences. That word classes exist satisfying both (1) and (2) exist has been shown by the linguists Beth Levin and R. M. W. Dixon. Literary style is measured by fitting GLMMs of word counts, where some of the factors used are author, time period, word choice, and type of grammatical structure. By analyzing both words and grammatical structures, it is hoped that both conscious and habitual patterns of an author can be detected. Finally, the GLMM fitting is done with both SAS and R, and the respective results are compared.

An All Configurations Approach To Testing For Latent Group-Based Interaction Effects In Two-Way Unreplicated Experiments.

✦Christopher Thomas Franck, Virginia Tech, 213 A Hutcheson Hall, Blacksburg, VA 24061, *chfranck@vt.edu*; Jason Osborne, North Carolina State University

Key Words: unreplicated, experiment, interaction, non-additivity

Statistical models which are used to analyze unreplicated factorial experiments typically do not include interaction effects between the experimental treatments. This is because the standard tests which determine whether or not interaction is present do not have an appropriate error term. Many authors have developed models and statistical tests based on restricted forms of interaction to address non-additivity for these unreplicated experiments. Our strategy is to instead assume that levels of one of the factors belong to some smaller number of latent groups. We develop an all configurations method to test for latent group-based interaction. Our model is able to capture latent group-based non-additivity since we allow groups to interact with the non-grouped treatment, but we assume that treatment effects are additive within group. Our all configurations method for detecting latent group based-interaction is shown to be more powerful than several methods for various forms of non-additivity.

Understanding Heterogeneity In Disease Progression In Patients With Primary Open-Angle Glaucoma

✦Feng Gao, Washington University School of Medicine in St. Louis, 660 S Euclid Ave, Division of Biostatistics, Box 8067, St. Louis, MO 63110 USA, *feng@wustl.edu*; J. Philip Miller, Washington University School of Medicine in St. Louis; Julia Beiser, Washington University School of Medicine in St. Louis; Chengjie Xiong, Department of Biostatistics, Washington University; Mae Gordon, Washington University School of Medicine in St. Louis

Key Words: Latent class growth model, multivariate longitudinal data, joint modeling

Primary open-angle glaucoma (POAG) is among the leading causes of blindness in the United States and worldwide. It is characterized by a gradual loss of vision and the progression rate is often described using linear regression models that are only focused on the average change over time. However, the progression of POAG is more likely to be heterogeneous and such an average effect could mask important individual differences. In this paper, we use latent class growth model (LCGM) to assess the heterogeneity of visual loss based on the 274 patients developed POAG from the Ocular Hypertension Treatment Study (OHTS). The visual deterioration as measured by mean deviation (MD) index and patterns standard deviation (PSD), two standard global summary measures of visual field test, was described using a piecewise linear function to allow different progression rates during pre- and post-POAG periods. The optimal latent classes, their probabilities, and the predictive baseline factors were estimated for each measure. Joint modeling of multiple longitudinal data was also performed to assess the co-trajectory of visual loss and structural change over time.

Comparison Of Longitudinal Models In Estimating The Rate Of Cognitive Decline Among Ad Patients

✦Yorghos Tripodis, Boston University, 801 Massachusetts Avenue, Department of Biostatistics, Boston, MA 02118, *yorghos@bu.edu*

Key Words: mixed-effect models, state-space models, Alzheimer's Disease

A number of statistical models has been used to model rate of decline among AD patients. We assessed different statistical methods in estimating effects of dichotomous variables on the rate of decline among AD patients from population and clinical studies. We generated 1000 perturbations where we randomly selected a subset that have a dichotomous characteristic that supposedly affects the rate of decline and we added it to each selected trajectory.We investigated different types of effects. We then fitted four different mixed effects models to allow for different patterns in the rate of decline: a) linear time and random intercept and slope, b) linear and quadratic time with random intercept and slope, c) linear and quadratic time with random intercept, slope and time squared d) state-space model with random walk plus drift. We calculated the bias and variance of the effect for each model.The choice of a statistical model affects the parameter estimates of the rate of decline and it has also an effect on its standard error. This has repercussions in studies where the rate of decline is the main outcome, as well as clinical studies that investigate disease-modifying effects.

Modeling Age At Onset And Progression Of Disability In Longitudinal Studies

✦Kumar B Rajan , Rush University Medical Center, 1645 W Jackson Blvd, Chicago, IL 60612, *kumar_rajan@rush.edu*

Key Words: Two-part model, survival and progression, disability study, longitudinal data analysis

Two-part models has been frequently used to study the onset and progression of severity in disability among older adults. However, the effect of risk factors on age at onset and progression are not well established. The focus of this study is to investigate the effect of cognitive and physical functions on the age at onset and the progression of disability after onset. Disability was assessed using the activities of daily living (ADL) measured annually over a 10-year period. The risk factors associated with the age at onset and progression of the disability were studied using a two-part latent model with a Weibull survival model for age at onset and a negative binomial model for trajectory of severity, with random effects from a jointly normal distribution. The marginal likelihood after integrating the random effects did not have a closed form; therefore, we integrated the joint likelihood using Gaussian-Hermite quadrature. The results showed that black males with low education, low cognitive and physical functions had a lower age at onset of disability. Additionally, cognitive and physical functions were negatively associated with the progression of severity after onset.

Power And Sample Size For Three-Level Cluster Designs

✦Tina Duong Cunningham, University of California Los Angeles, 827 Levering Ave, Apt # 801, Los Angeles, CA 90024, *cunninghamdt@ucla.edu*; Robert E. Johnson, Virginia Commonwealth University

Key Words: Sample Sizes, Cluster Randomized Trials, Three-level design

In designing studies with multilevel structure, one of the most important tasks is to implement proper statistical methods in computing power and sample sizes. The evaluation of required sample size for three-level designs is complicated because of several reasons. First, the multilevel nesting effect introduces more than one intracluster correlation into the model. Second, the variance structure of the estimated treatment difference is more complicated. Third, sample size computations for several levels are required. In this work, we developed sample size and power formulas for the three-level data structures based on the generalized linear mixed model approach. We derived power and sample size equations for detecting a hypothesized effect. We developed a SAS user-interface macro that allows the researcher to estimate sample size for three-level design for different scenarios. The formulas take into account the fact that random assignment can take place in different levels of the study design.

574 Advanced Methods in Health Outcomes with Policy Relevance

Section on Quality and Productivity, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Causal Mediation Analysis For Nonlinear Models With Confounding

✦Jeffrey M. Albert, Case Western Reserve University, Department of Epidemiology and Biostatistics, 10900 Euclid Avenue, Cleveland, OH 44106-4945, *jma13@case.edu*

Key Words: empirical distribution function, generalized linear model, IPTW

Estimation of direct and indirect (or mediation) effects is conventionally performed using a product of coefficients approach based on multiple regression or structural equations models. In the case of nonlinear models, causally interpretable estimates of mediation effects may require a nonstandard expression, referred to by Pearl (2011) as the mediation formula. Implementation of the mediation formula typically uses an assumed probability distribution for the mediator based on a parametric model. We develop an approach that avoids such distributional assumptions, by using the empirical distribution function for the mediator. A challenge that arises when applying this method to observational data, is to adjust for measured confounders of the exposure-mediator relationship. We consider weighting approaches to estimate causal mediation effects while adjusting for confounding in the above context. Alternative approaches are compared via simulation studies. The methods are applied to data from a study of the effect of very low birth weight on dental caries in adolescence.

The Peters-Belson Method For Assessing Health-Care Disparities, Extended To Survival Analysis

Lynn Eberly, University of Minnesota; ✦James S. Hodges, Div of Biostatistics, U of Minnesota, 2221 University Ave SE, Ste 200, Minneapolis, MN 55455, *hodge003@umn.edu*; Donna Z. Bliss, School of Nursing, U of Minnesota; Kay Savik, School of Nursing, U of Minnesota; Olga Gurvich, School of Nursing, U of Minnesota; Susan L. Harms, College of Pharmacy, U of Minnesota; Christine A. Mueller, School of Nursing, U of Minnesota

Key Words: disparities, survival analysis, Peters-Belson method

The Peters-Belson method was developed for quantifying and testing disparities between groups, using linear regression to compute group-specific observed and expected outcomes. It has since been extended to logistic regression for binary outcomes. For an NIH-funded project assessing racial/ethnic disparities in nursing home care (the REDSKIN Study), we extended the Peters-Belson approach to survival analysis, including stratified analyses. The extension uses the theory and methods of expected survival based on Cox regression in a reference population, as developed in Therneau & Grambsch (Modeling Survival Data, 2000; Springer; Chapter 10). We used the survival package in the R system to do computations; Therneau & Grambsch give a SAS macro that could be adapted readily. We describe the extension, show how we applied it in the REDSKIN Study, and discuss some issues in implementing it.

Unknown Correlation Coefficient Between A Primary And A Secondary Endpoint In A Two-Stage Group Sequential Design

✦Yi Wu, Northwestern University, Dept of Statistics, 2006 Sheridan Rd, Evanston, IL 60208, *yi.ripple@gmail.com*; Ajit C Tamhane, Northwestern University; Cyrus Mehta, Cytel Inc.

Key Words: group sequential design, clinical trials, multiple endpoints, correlation coefficient, familywise error rate

We consider a two-stage group sequential design with a primary and a secondary endpoint where the secondary endpoint is tested only if the primary endpoint shows significance. We assume that the two endpoints follow a bivariate normal distribution with unknown correlation coefficient \rho. Setting \rho=1 will provide the most conservative critical boundary (Tamhane, Mehta and Liu 2010). However, replacing \rho with its sample estimate r will cause familywise error rate (FWER) inflation when r is less than \rho. We have developed an approach to control the overall FWER with adjusted second stage critical boundary utilizing the upper confidence limit of \rho. The power gain for rejecting the secondary hypothesis of the new approach over the conservative method is discussed.

Model Selection Stability For Population Health Risk Models

✦Diane M. Richardson, Center for Health Equity Research and Promotion, Philadelphia VA Medical Center, 3900 Woodland Avenue, Bldg 4100, Philadelphia, PA 19104 USA, *dianemrichardson@gmail.com*; Daniel Louis, Thomas Jefferson University; Mary Robeson, Jefferson Medical College

Key Words: risk model, selection, stability, population health, hospitalization

Clinical diagnostic data available in regional health databases is a rich source of information for modeling population health outcomes. Models that predict patient risk are used for case-mix adjustment, profiling, and resource allocation. However, selection of a meaningful subset of predictors is plagued by instabilities that limit the practical utility of the models. To assess the potential of newer methods not widely used in population health applications to improve model stability, we developed models for hospitalization risk in a general population, comparing stepwise, lasso, and stability-based selection. We used health system data to create over 200 clinical and disease stage indicators as predictors, and modeled future risk of potentially avoidable hospitalization. We report measures of model selection and prediction stability, and illustrate issues relevant to the selection of highest risk patients for enrollment in enhanced care management programs. Our results demonstrate that model selection stability can be improved markedly, and suggest that model selection stability be reported as a performance measure for population health risk models.

Accounting For Skewness With A Nonparametric Estimate Of The Incremental Cost Effectiveness Ratio

✦David K. Blough, University of Washington, Department of Pharmacy, Box 357630, Seattle, WA 98195 United States, *dkblough@u.washington.edu*

Key Words: bivariate location, skewness, incremental cost effectiveness ratio

Typically, mean costs and mean measures of effectiveness are used in the estimation of the incremental cost effectiveness ratio (ICER). However, cost and effectiveness distributions are often skewed. The goal of this work is to use a nonparametric bivariate generalization of location to obtain a location region in the cost effectiveness plane. Analogous to Mann-Whitney, all pairwise cost and effectiveness treatment differences are computed. Using univariate methods, an estimate of a location rectangle is obtained as the Cartesian product of marginal location intervals, intervals that account for skewness. By considering the intersection of such rectangles over all orthogonal transformations, a closed convex region is obtained. Kolmogorov-Smirnov provides simultaneous confidence bands for each bivariate rectangle. Any point of this region is an estimate of the ICER; all such points satisfy basic axioms of bivariate location. Comparing the location region with the usual bootstrap estimates for the ICER, it was found that the location region approach produced confidence intervals that were considerably larger for both skewed and symmetric distributions of costs and effectiveness.

Modeling Zero- Inflated Longitudinal Ordinal Data: An Application To Marijuana And Cocaine Use

✦Rajendra Kadel, University of South Florida, , *rkadel@mail.usf.edu*; Getachew Dagne, University of South Florida

Key Words: Zero- Inflated, Ordinal Response, Bivariate Probit, Longitudinal Data, Drug Abuse

Ordinal response data such as severity of pain, degree of disability, satisfaction with healthcare provider are prevalent in many areas of research including biomedical and social science research. Besides the ordinal nature of the data, there are often excess zeros occurring when measuring abnormal behavior such as illicit drug use, symptom and side effect of rare diseases. For example, in the Drug Abuse and Treatment Outcome Study (DATOS), where a cohort of adolescents were followed prospectively to study the effectiveness of adolescent drug treatment, the outcome data are ordinal data with excess zeros. Zero inflation in ordinal categorical data coupled with longitudinal structure makes it difficult to analyze and interpret. The traditional multinomial logit or probit model is not very appropriate to analyze such data. We propose a zero-inflated bivariate ordinal probit model to analyze longitudinal ordered categorical data with excess zeros. The methods proposed are illustrated with a real application of marijuana and Cocaine use data from DATOS, 1993-1995 and the results are compared with the existing methods.

575 Reliability Methods: Censoring, Estimation and Control

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

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Two Robust Estimation Techniques For Monitoring Reliability

✦Derya Karagoz (Caliskan), Hacettepe University, Hacettepe University, Department of Statistics, Beytepe, Ankara, International 06800 Turkey, *deryacal@hacettepe.edu.tr*

Key Words: Control charts, Moment approximation, Repeated median estimator, Median/Qn estimator

In this study, we propose robust methods to monitor reliability by using control charts when the distribution of inter-failure times is two-parameter Weibull. Control charts are useful for monitoring reliability. To estimate the unknown parameters we use explicit and robust estimation techniques which are based on repeated median and median/Qn estimators. We are interested in not only inter failure times, but also the cumulative time elapsed between failures (Tr). Since we do not know the distribution of the sum of independent Weibull random variables , we propose two approximation methods. Two -moment normal approximation and three- moment chi-square approximation methods are applied in order to obtain approximate confidence interval limits. Moment approximations provide reasonably accurate results to the percentage points of distribution Tr .

A Special Nonhomogeneous Poisson Process Estimation For Window-Observation Repairable System

✦Ming Li, Applied Statistics Lab, GE Global Research, 1 Research Circle, Niskayuna, NY 12309, *limi@ge.com*; Brock Osborn, Applied Statistics Lab, GE Global Research ; Yili Hong, Department of Statistics, Virginia Tech

Key Words: Reliability, Repairable System, Nonhomogeneous Poisson Process, Window-Observation Data, Forecast

A fleet of heavy-duty industry equipment is used in infrastructure constructions. The age of the equipment usually exceeds 20 years, and during its useful life repair and maintenance are needed. Unfortunately, not all repairs or maintenance events are recorded. The historical data set consists of series of observation windows in which the equipment is monitored and events are recorded. The time between observation windows varies from several days to several years. To model the reliability of the equipment and to forecast the cost of future repairs and maintenance events, a special three-piece nonhomogeneous Poisson process method is applied to the window-observation data. The bath-tub intensity results provide useful information for reliability estimations.

Statistical Inference Of Adaptive Progressively Censored Data With Lognormal Lifetimes

✦Fang Duan, Southern Methodist University, Dallas, TX 75275, *fduan@smu.edu*; Hon Keung Tony Ng, Southern Methodist University

Key Words: Life-testing, Type II progressive censoring, maximum likelihood estimator, least square estimation

An adaptive progressive censoring scheme, which is useful in life-testing or reliability experiments, is introduced in this talk. This censoring scheme can be viewed as an efficient design in which the number of observed failure times is pre-fixed, and the total time on test will not be too far away from an ideal time limit. Parameter estimation for the lognormal distribution is discussed. Different estimation methods, including maximum likelihood estimation (MLE), approximate MLE and least square estimation are compared using Monte Carlo simulation.

Service Life Prediction Using Accelerated Degradation Data From Laboratory Testing And Outdoor Weathering Data

✦Yili Hong, Department of Statistics, Virginia Tech, 213 Hutcheson Hall, Virginia Tech, Blacksburg, VA 24061 USA, *yilihong@vt.edu*; William Q. Meeker, Iowa State University

Key Words: Coatings, Photodegradation, Reliability, UV exposure

Photodegradation caused by ultraviolet (UV) radiation is a primary cause of failure for coatings, as well as many other products made from organic materials exposed to sunlight. Other environmental factors includes temperature and humidity. Service life prediction of such coatings is difficult because UV radiation from the sun are highly variable. Both indoor and outdoor experiments were conducted at the US NIST to generate necessary experimental data. Accelerated laboratory tests were conducted on a device in which spectral UV wavelength and intensity, temperature, and relative humidity (RH) can be precisely and accurately controlled over time. Outdoor exposure experiments were conducted on the roof of a NIST laboratory. The chemical degradation of coatings was measured every few days. Longitudinal information on temperature, RH and solar spectrum for outdoor specimens were also recorded. In this paper, a physically motivated model is fitted to indoor data. Outdoor data were used to validate the model fitted to the indoor data. Service life prediction was obtained based on the fitted model and prediction intervals were obtained to quantify the statistical uncertainties.

The Cost Of Reliability: Demonstrating The Financial Benefit Of Reliability Testing

✦Robert O'Donnell, Hewlett-Packard, 1000 NE Circle Blvd, MS 323B, Corvallis, OR 97330, *bob.odonnell2@hp.com*

Key Words: reliability, MonteCarlo, cost, test

The implementation of lean strategies in industry has driven the need to more directly justify the cost of reliability testing. Testing can provide benefits (reduced warranty costs and increased customer satisfaction) but also involves direct costs. Reliability engineers look for ways to demonstrate the financial benefits of product testing by modeling testing as a decision making process. It is a decision making process with uncertainty in the true reliability of the tested product and the costs involved in rectifying problems found; and choices to be made regarding sample size and confidence level. A Monte Carlo simulation based procedure will be presented for demonstrating at a product level the expected financial cost of reduced product reliability. In addition, methods will be described for assigning an economic value to each of a number of reliability tests (targeting different failure modes). These methods can be used to identify an economically optimal sample size for a test as well as comparing the benefit of that optimal test to the benefit (or cost) of not doing the test at all.

A New Exponential Goodness-Of-Fit Test For Data Subject To Ordinary And Multiply Type Ii Censoring

Scott Lesch, Riverside Public Utilities; ✦Daniel R. Jeske, University of California, Riverside, Department of Statistics, 1340 Olmsted Hall, 900 University Ave., Riverside, CA 92507, *danel.jeske@ucr.edu*

Key Words: Goodness-of-fit, Order statistics, Quadratic forms

We begin by reviewing multiply type II censoring and give examples as to how it can arise in applications. We then present a new goodness-of-fit (GOF) test statistic for both ordinary and multiply Type II censored Exponential data. This test statistic is based on a ratio of linear functions of order statistics. We discuss results from a power study that shows the test compares favorably to currently available GOF tests.

Generalized Linear Modeling For Assessment Of A Performance Based Logistics Strategy (Pbl) On Demand/Cost Reductions In An Aging Complex System

✦Mark Carpenter, Auburn University, 221 Parker Hall, Department of mathematics and statistics, Auburn, AL 36849-5310, *carpedm@auburn.edu*; Wesley Randall, Auburn University

Key Words: discrete valued time series, non-homogeneous Poisson, mixture modeling, Bayesian analysis, performance based logistics

This paper provides a methodology to investigate a hypothesized link between spare part consumption, reliability improvement, cost reductions, increases in mean time between failures (MTBF), and a Performance Based Logistics strategy. More specifically we investigate the assertion that a 5 year firm fixed price (FFP) PBL contract created an incentive for prime contractor to invest in repair process and reliability improvements, which resulted in specific cost reductions through efficiency gains and cost reduction in maintenance, repair, and overhaul (MRO) type activities. Assuming that the observed demands for each part or collection of parts over time are from a non-homogeneous Poisson process, we investigate short- and long-run trends in demand over time, using generalized Poisson regression, modeling time series of counts (discrete valued time series), Bayesian models, hierarchical and discrete mixture modeling.

576 Generalized Linear and Mixture Models

Section on Statistical Computing, Section on Statistical Graphics, Section for Statistical Programmers and Analysts, Section on Statistics and the Environment

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Biometric Face Recognition And Performance Evaluation

✦Geof H Givens, Colorado State Univ., Dept. of Statistics, CSU, 1877 Campus Delivery, Fort Collins, CO 80523, *geof@lamar.colostate.edu*; Ross Beveridge, Colorado State Univ.; Jonathon Phillips, NIST; Bruce Draper, Colorado State Univ.; Yui Man Lui, Colorado State Univ.; David Bolme, Colorado State Univ.

Key Words: biometrics, face recognition, GLMM

The science of biometric face recognition is relatively unfamiliar to statisticians. We describe the major aspects of the field and outline several recognition algorithms ranging from an elementary principal components approach to quite sophisticated methods. Evaluating algorithm performance and how it relates to human, image, and environmental variables is key to advancement of the field. We summarize one performance study using a state-of-the-art commercial recognition algorithm and a generalized linear mixed model to illustrate the types of conclusions that can be made to better understand recognition performance, improve real-world applications, and develop innovative new face recognition algorithms.

Iteratively Reweighted Poisson Regression For Fitting Generalized Linear Model With Multiple Responses

✦Yiwen Zhang, North Carolina State University, 2311 Stinson Drive, Raleigh, NC 27695, *yzhang31@ncsu.edu*; Hua Zhou, North Carolina State University

Key Words: Dirichlet-multinomial, GLM, MM algorithm, multinomial-logit, multiple responses, negative-multinomial

Generalized linear models with multiple responses (MGLMs) are seeing wider use in modern applications such as pattern recognition, document clustering, and image reconstruction. Examples of MGLMs include multinomial-logit models, Dirichlet-multinomial overdispersion models, and negative-multinomial models. Maximum likelihood estimation of MGLMs is difficult due to the high-dimensionality of the parameter space and possible non-concavity of the log-likelihood function. In this article, we propose iteratively reweighted Poisson regression as a unified framework for maximum likelihood estimation of MGLMs. The derivation hinges on the minorization-maximization (MM) principle which generalizes the celebrated expectation-maximization (EM) algorithm. MM algorithm operates by constructing a surrogate function with parameters separated. Optimizing such a surrogate function drives the objective function in the correct direction. This leads to a stable algorithm which possesses good global convergence property and is extremely simple to code. The proposed algorithm is tested on classical and modern examples.

Geometry Of Generalized Linear Models

✦George R Terrell, Statistics Department, Virginia Polytechnic Institute, Blacksburg, VA 24061, *terrell@vt.edu*

Key Words: Loglinear models, logistic regression, convex regression, duality

It has long been found useful to think of classical linear regression geometrically, as involving predictions and errors that lie in certain linear subspaces of observation space. Generalized linear models, such as logistic regression, by contrast, are usually formulated in terms of their likelihood. We will here show that a rich class of models, formulated in terms of vector geometry in observation space, includes the generalized linear models with canonical link function. Their geometry is formally dual to the classical case. The new characterization simplifies computation in a number of cases.

Robust Estimation Of Censored Mixture Models

✦A.M. Santos, SimpleGeo and UC-Denver, , *pintyesantos@gmail.com*; Karen Kafadar, Indiana University

Key Words: robustness, finite mixtures, EM algorithm, biweight

This research is motivated by the common problems of censoring and mixtures in real-world data. Measurement limitations cause many datasets in healthcare, economics, business and other research areas to have inexact results at the upper and lower bounds of the data. Additionally, many datasets contain multiple populations, resulting in mixture distributions. We demonstrate both a MLE method for fitting this type of data and a robust extension with comparable estimation properties and improved computational time. We evaluate both methods on multiple scenarios, including the possibility that the underlying distributions are not Gaussian, but have heavier tails.

Single Molecule Super-Resolution Analysis: Using Mixture Models To Accurately Localize And Enumerate Fc$\Varepsilon$Ri Receptors

✦W. Duncan Wadsworth, University of New Mexico, Department of Mathematics and Statistics, 1524 Sigma Chi NE, Albuquerque, NM 87106, *dwadswor@unm.edu*; Keith Lidke, University of New Mexico, Department of Physics and Astronomy

Key Words: Mixture Models, EM Algorithm, Image Analysis, Statistical Computing with GPUs, Fluorescence Spectroscopy

Due to the important role of membrane proteins in cell signalling, accurate localization and enumeration of molecules on a cell surface is of great importance to theoretical and experimental biology. Developments in fluorescence microscopy allow super-resolution (SR) image localization of individual fluorophores with a precision near 20 nm. However, using a set of observations to localize and enumerate proteins is complicated by finite localization precision and miscounting due to blinking, under-labeling and photo-bleaching. We describe the use of mixture models to analyze single-molecule SR data resulting in accurate estimates of both fluorophore location and number. The Expectation-Maximization (EM) algorithm parameterizes the mixture model, where component modes represent estimated fluorophore locations. We include a-priori knowledge of labeling efficiency and fluorophore characteristics. An implementation of the EM algorithm using Nvidia CUDA code runs on Graphics Processing Units, significantly decreasing computing time. Results are presented for simulated data and experimental data acquired using fluorescent dyes bound to the FcεRI receptor.

577 Nonparametric Methods for Complex Dependence ●

Section on Nonparametric Statistics, International Indian Statistical Association

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Empirical Likelihood Confidence Intervals For Complex Survey Data

✦Nancy L Glenn, Ph. D., Texas Southern University, 874 Yorkchester Dr. #121, Houston, TX 77079, *nglenn@alumni.rice.edu*; Andrea Shelton, Texas Southern University; Andrea Shelton, Texas Southern University

Key Words: nonparametric, empirical likelihood, complex sample survey data

Complex sample surveys which contain a large representative sample of various demographic groups provide excellent sources of data for accessing health and nutrition. Data from sample surveys are obtained through multistage sampling designs that involve clustering, stratification, and nonresponse adjustments. Using standard statistical methods in this context induces a nonstandard covariance structure among sample quantities. Standard statistical methods are typically not applicable to sample surveys because independent, identically distributed observations seldom result from such complex sample surveys. We derive empirical likelihood confidence intervals for complex survey data. Using data from the Third National Health and Nutrition Examination Survey we construct empirical likelihood confidence intervals.

Multivariate Linear L1 Regression For Cluster-Correlated Data

✦JAAKKO NEVALAINEN, University of Turku, Statistics, Assistentinkatu 7, Turku, International 20014 Finland, *jaakko.nevalainen@utu.fi*; KLAUS NORDHAUSEN, University of Tampere; HANNU OJA, University of Tampere

Key Words: clustered data, L1 objective function, spatial sign

We consider the multivariate linear regression model with a p-variate response variable, q-variate vector of explanatory variables and a p-variate random error. The goal is to make inference on the unknown q x p regression coefficient matrix. Commonly, the estimation of the parameters is based on L2 or L1 objective functions. However, the standard assumption of the independence of the random errors does not hold if the data are clustered; they are correlated. In this talk we review the multivariate L1 regression theory in the case of iid error variables, and then we extend the asymptotic theory to the cluster-correlated case, including weighted L1 estimates of regression coefficients. The theory is illustrated with data examples and a simulation study.

Local Polynomial Regression And Mixed Models

✦Li-Shan Huang, Institute of Statistics, National Tsing Hua University, HsinChu, 300 TAIWAN, *lhuang@stat.nthu.edu.tw*; Kung-Sik Chan, Dept of Statistics, University of Iowa

Key Words: Trigonometric basis functions, projection, ANOVA decomposition

Penalized spline smoothing has been shown to connect to linear mixed models in statistical literature. In this paper, we show that local polynomial regression when expressed in a projection framework also has interesting connections to mixed models. Under some conditions, the local polynomial projection approach admits an equivalent mixed model formulation where the fixed-effects part includes the polynomial basis functions. Then we further show that asymptotically the trigonometric functions are the penalized basis in the mixed model

formulation. These results suggest a new smoothing approach using a combination of unpenalized polynomials and penalized trigonometric functions. We illustrate the potential usefulness of the new approach with real data analysis.

Linear Latent Structure Analysis (Lls)

✦Mikhail Kovtun, Duke University, Durham, NC 27708, *mikhail.kovtun@duke.edu*; Igor Akushevich, Duke University; Anatoliy Yashin, Duke University

Key Words: mixed distributions, latent structure, identifiability, mixture invariants, estimation, consistency

LLS analysis is aimed to model a joint distribution of a number of categorical random variables (survey-like data) as a mixture of independent distributions. Generally, the mixing distribution cannot be identified. However, if the mixing distribution is carried by a low-dimensional subspace of the space of independent distributions (which is when LLS analysis is applicable), the carrying subspace of the mixing distribution and a number of low-order moments of the mixing distribution can be identified. The number of identifiable moments increases as the number of categorical random variables increases, and in the case of infinite number of random variables the mixing distribution is identifiable. Under modest assumptions, the identifiable invariants of the mixing distribution can be consistently estimated, and there exists an efficient computational algorithm for estimating identifiable invariants.

Probabilistic Index Mixed Models For Clustered Data

✦Fanghong zhang, Ghent University, Belgium, Krijgslaan 281, Campus de sterr, University Gent, Ghent, International 9000 Belgium, *fanghong.zhang@ugent.be*

Key Words: semi-parametric inference, probabilistic index, pseudo likelihood estimation, mixed models, rank statistics

When the data distribution is skewed and cannot easily be made symmetric by transformations, routine parametric models can be problematic. In those cases probabilistic index models can be successfully applied (Thas et al,2010). Probabilistic index models are semi-parametric regression models for the probabilistic index Pr(Y < Y*) in function of corresponding covariate values X and X* (Thas et al,2009). Here, we extend these models to enable the analysis of clustered data. Specifically, given two random records (Y_ik,X_ik) and (Y_jl,X_jl), where the k/l outcome Y is from subject i/j, models of the form: logit[P(Y_ik< =Y_jl | X_ik,X_jl,b_i,b_j)] = (X_ik-X_jl)*beta +(b_i-b_j). Random effects b_i(j) are assumed to be mutually independent and normally distributed with mean zero and constant variance. Consistent asymptotically normal estimators are established through a semi-parametric procedure involving pseudo-likelihood estimation. Estimators can be used to construct rank-based non-parametric tests for clustered data which enable adjustment for covariates. Simulation studies evaluate the finite-sample behavior of estimators and data analysis results will be provided.

Sequential Statistical Inverse Problems

✦Darren Warren Homrighausen, Department of Statistics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, *dhomrigh@stat.cmu.edu*

Key Words: inverse, ill-posed, repeated, statistical

In many signal processing applications, the same signal is recorded many times. In some cases, as in telescopic imaging and MRI, the signal can only be recorded indirectly and under some noise, both of which are changing over time. This regime is different than the more common inverse problem formulation where theoretical analysis centers on an in-fill asymptotic. In the sequential regime, the resolution on the sensor does not improve over time and hence the only relevant asymptotic is with sample size. In this talk, I propose both new statistical methodology to the sequential setting and analyze previously used methods in a more rigorous framework.

Population Intervention Causal Effects Based On Stochastic Interventions

✦Ivan Diaz Munoz, UC Berkeley, 114 Haviland Hall, UC Berkeley, Berkeley, CA 94709, *ildiazm@berkeley.edu*; Mark van der Laan, University of California, Berkeley

Key Words: causal effect, stochastic intervention, TMLE

The causal effect of an intervention on a population of interest typically concerns interventions that deterministically assign the treatment or exposure in a static or dynamic way. We propose a new causal parameter that takes into account the fact that intervention policies can result in stochastically assigned exposures. The statistical parameter that identifies the causal parameter of interest is established, and a double robust Targeted Maximum Likelihood Estimator (van der Laan and Rubin, 2006) is developed. A simulation study is performed to demonstrate the double robustness of the TMLE, and an application example using physical activity data is presented.

578 Analysis and Modeling of Networks

■●

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

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Optimal Path On Network With Stochastic Flow

✦WANLI MIN, IBM Singapore, , *minw@sg.ibm.com*

Key Words: Gaussian Process, Variation Method, Heteroscedasticity

We consider the problem of optimal path between origin and destination over a network where each segment of the path incurs certain time varying cost and such cost may well follow some probabilistic distributions. We construct a differential equation to solve the problem in its elementary form: given the cost across segments of the network at any future time point. We introduce a variation technique to establish a general representation in a form of the elementary solution, allowing the cost over network to be characterized by certain Gaussian field. Finally, we apply the results to a problem of fastest driving route giving future traffic conditions on road network.

Statistical Inferences For The Offered Optical Network Unit Load

✦Sumith Gunasekera, The University of Tennessee at Chattanooga, Department of Mathematics, 615 McCallie Avenue, Dept. 6956, Chattanooga, TN 37403-2598, *Sumith-Gunasekera@utc.edu*

Key Words: Classical tests, Generalized tests, Pareto distribution, Offered Optical Network Unit Load, Multiplexing

A Passive Optical Networks (PONs) system of multiple sources that generates sequences of ON-and OFF-periods is considered. When multiple parallel connections are aggregated (serialized or multiplexed), the net load, called the Offered Optical Network Unit Load (OOL), is generated at the Optical Network Unit (ONU). Statistical inferences of OOL of Self-Similar (SS) or Long-Range-Dependent (LRD) network traffic are performed. An approximate test and approximate confidence intervals based on the Large Sample Normal Approach as well as a generalized variable test and generalized confidence intervals based on the Generalized Variable Method (GVM) are given. A limited simulation study is given to demonstrate the advantages of the propopesd Generalized Variable Method over the Approximate Procedures.

Availability Analysis Of A Large-Scale Networked System

✦Shuguang Song, The Boeing Company, Boeing Research & Technology, P.O. Box 3707, MS: 14-27, Seattle, WA 98124, *shuguang.song@boeing.com*

Key Words: availability, networked system, spatio-temporal, failure time, repair time, maintenance

In modern industry, complex networked systems have become more and more common. The availability of complex systems is an active research area with application not only in engineering but also in experimental and social sciences. Consider a large-scale networked system with possible heterogeneous N1 sensors, N2 back-up units, i.e. failure time and repair time of different sensors/back-up units could follow different distributions. If a sensor fails at a site, a back-up unit will be transported to the failed sensor site and replace the failed sensor if there is at least one available back-up unit before the failed sensor is repaired back to operating condition. Once the failed sensor is back in operation, then the back-up unit will be immediately replaced by the repaired sensor. Due to the design of the networked system, the physical capability of a sensor/back-up unit, and terrain variations and weather conditions across a large-scale networked system, the "working availability" of the large-scale networked system may not only be time indexed, but also be space indexed. In this paper, we present some results about the availability analysis of such a large-scale networked system.

Classification Methods For Predicting Process Run Time In Computing Systems

✦Wei Chen, Operations Research and Information Engineering, Cornell University, 206 Rhodes Hall, Cornell University, Ithaca, NY 14853, *wc438@cornell.edu*; Dawn Woodard, Operations Research and Information Engineering, Cornell University; Moises Goldszmidt, Microsoft Research Silicon Valley

Key Words: time series, classifier, functional data analysis

In distributed computing systems, multiple servers perform similar tasks simultaneously. Although the tasks have similar characteristics, the running times between servers may vary widely, for example due to data skew, servers competing for data access, and problems with the servers. Early prediction of abnormally slow performance can be used in system management; for instance one can restart the same task on a different server (closer to the data) to save total running time, or probe the server to verify it is working properly. We introduce statistical methods for predicting whether servers in a computing system (Microsoft's DryadLINQ) will perform abnormally slowly, based on system data early in the process. A natural approach is given by standard functional data analysis techniques; however, these models do not capture an important feature of the server data: regime-switching behavior, where the transitions between the regimes occur at random times. We introduce several time series models that allow this regime-change behavior and relate it to the outcome variable (process completion time). We describe efficient computational methods, including a blocked Gibbs sampler.

A Statistical Investigation Of Swamp/Network Dynamics

✦Morris Morgan, Hampton University, Queen and Tyler Streets, Hampton, VA , *morris.morgan@hamptonu.edu*; Carolyn Bradshaw Morgan, Hampton University

Key Words: Robustness, Simulations, Normal Distribution, Networks

The current research is focused on assessing the statistical robustness of synchronized network designs used to describe swamp behavior. Guidelines are developed for determining the necessary number of controllers (linkages) needed to maintain stable swamp behavior and the statistical sensitivity of such global systems to breakdowns induced via internal dynamic fluctuations arising from the decay of the controller connections. It is shown that the ability to predict long-term swamp synchronization correlates well with an initial spatial metric estimated at the start of the dynamic simulation process. Both the MATLAB Simulink and SPlus software packages where used to conduct the statistical investigations.

Testing For Electrical Leakage: A Bayes/Empirical Bayes Approach

✦Cheng Chung Chen, Math Dept, Texas A&M University at Kingsville, &00 University Blvd,, Kingsville, TX 78363 USA, *kfccc01@tamuk.edu*; Hongbin Chen, Academia Sinica

Key Words: Circuit Testing, Group Testing, Bayes Estimation, Empirical Bayes, Networks

Abstract this paper considers the problem of detecting and locating electrical leakage between two groups of nets. This problem was previously considered by Skilling (1982) and a clever method was patented by him with complexity of O(n). Chen and Hwang (1989) improved the complexity to O(2log2N) by using group testing results. Notice that both of the above two approaches are deterministic in nature. Furthermore, it was observed that only a very small subset of networks can be close enough to possibly short together during the manufacturing process as evidenced by failure data. The current study uses the Bayes /Empirical Bayes techniques to capture the topological structure of the network and defect data base adaptively. We The proposed procedure compares favorably with both Chen& Hwang and Skilling's method. The complexity is bounded above by 2l?og?_2?n. Furthermore, is also consider the case in which one of the two groups of nets being tested is restricted in size, as is true in most practical applications.

Model Development Of Genetic-Based Algorithm To Solve The Capacitor Placement Problem In Primary Feeders

✦Mohammad Abul Quasem, Howard University, Dept. of Information Systems & Decision Sciences, School of Business, 2600 6TH STREET, Washington, DC 20059, *mquasem@howard.edu*; Mohamed H. Shwehdi, King Fahd University of Petroleum & Minerals, Electrical Engineering Department

Key Words: Capacitor Placement, Distribution Feeders, Power Distribution Systems, Repeated Iterations, Economics, Genetic-Algorithm

Nowadays fast growth in electricity consumption is one of the most challenging problems for the electric utilities. VOLT/VAR and losses at the distribution systems level are the key solution means for this problem. A solution to this problem can assist in offsetting the cost of energy and generating capacity expenses. Two new formulations of Capacitor Placement Problem (CPP) are proposed. The first one is based on capacitors/losses cost balance principle while the second one is based on a comprehensive cost evaluation of the network performance post CP. A genetic-based algorithm is developed and implemented to solve CPP. This algorithm is a process of repeated iterations, a continual search for the expected model of multiple load models to be compensated and then compensation is applied to it. When no compensation is needed for all load models, the iterations stop. This solution scheme is tested on experimental systems. The impact of capacitor installation on fault level and network resonance if also investigated.

579 Measurement Error and Data Quality

■

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Assessment Of Measurement Error And Nonresponse Error Using Respondent-Provided Paradata From Paper Self-Administered Questionnaires

✦Heather M. Schroeder, University of Michigan, SRC-SRO, PO Box 1248, G384A, Ann Arbor, MI 48106-1248, *schroedh@umich.edu*; Russell W. Stark, University of Michigan; Ashley Bowers, University of Michigan; Roger D. Goddard, Texas A&M University; Robin Tepper Jacob, University of Michigan

Key Words: paradata, measurement error, nonresponse error, paper self-administered questionnaires

Paradata analysis was introduced as a powerful tool to reduce error in web surveys and interviewer-administered surveys (Kreuter, 2010), and researchers are now harnessing the power of paradata analysis in paper self-administered surveys as well. Features of the environment in which self-administered surveys are completed can be measured using respondent-provided paradata. Interview setting features such as the presence of others have been shown to affect reporting in interviewer-administered surveys (see Tourangeau & Yan, 2007, for a review). We also expect that the survey setting may introduce measurement variability in paper self-administered surveys. Paradata including the presence of others, consultation with others in answering the questions, and location(s) of completion were collected from teachers and principals in selected schools as part of a self-administered survey. We assess whether features of the survey setting contribute to measurement error and nonresponse error as indicated by variation in the reporting of principals' and teachers' effectiveness, straightlining, reliability of multi-item scales, and missing data rates.

Designing Estimators Of Nonsampling Errors In Components Of Census Coverage Error

✦Mary H. Mulry, U.S. Census Bureau, 2884 Manorwood Trail, Fort Worth, TX 76109, *mary.h.mulry@census.gov*; Bruce D. Spencer, Northwestern University

Key Words: census omissions, census erroneous enumerations, net census coverage, nonsampling error

The 2010 Census Coverage Measurement Program (CCM) plans to use logistic regression models for the correct enumeration rate, data-defined rate, and match rate that will be used in dual system estimation of the population size which then will be used to measure net census coverage error. Recent studies of the error structure of the logistic regression estimator for net error have described how various kinds of sampling and non-sampling errors affect the estimates of the net error as well as estimates of omissions and erroneous enumerations. In addition, the studies have provided decompositions of the sampling and non-sampling errors. These decompositions have been useful in designing a schematic plan for using a simulation methodology to synthesize the effect of the sources of error on the estimates of net coverage error as well as on the estimates of omissions and erroneous enumerations. The sufficient statistics that will facilitate the simulation of the effect of errors have been identified. This paper discusses the design estimators of nonsampling errors suitable for use with data sources that will be available from the CCM and the evaluations of the CCM.

Using Interviewer Observations Related To The Interview Process For Measurement Error And Nonresponse Error Adjustment

✦Rachel Anne Orlowski, Wayne State University, Room 2207 AAB, 5700 Cass Avenue, Detroit, MI 48202 USA, *rachel.orlowski@wayne.edu*; Ashley Bowers, University of Michigan; Wayne E Baker, University of Michigan

Key Words: measurement error, nonresponse error, interviewer observations, paradata

Interviewer estimates of household characteristics have been used for nonresponse error adjustment (Kreuter et al., 2010) but rarely have interviewer observations of respondent behavior and attitudes related to the interview process been exploited for survey error adjustment (Tarnai & Paxson, 2005). In addition to adjustment purposes, these paradata also can guide interviewer training efforts and case assignment based on interviewer skills. In panel designs, these observations from prior waves can be used for adjustment and management in future waves. We analyze interviewer assessments of respondent understanding of the questions and respondent attitude toward the interview from a rotating panel telephone survey with a battery of attitudinal items. Indicators of measurement error, including selection of the same response for most of a question series, socially desirable reporting, and reliability of multi-item scales, are modeled as a function of interviewer assessments of respondent understanding and attitude. We also examine whether interviewer assessments can explain indicators of nonresponse error, such as missing data rates and participation in a future survey wave.

Assessing And Adjusting For Response Error Using A Multi-Phase Survey Approach In The Residential Energy Consumption Survey

✦Marilyn Worthy, Energy Information Administration, 1000 Independence Ave, SW, Washington, DC 20585, *marilyn.worthy@eia.gov*; Katie Joseph, Energy Information Administration

Key Words: multiphase survey, nonresponse bias, response error, household survey, RECS

The Residential Energy Consumption Survey (RECS) is a quadrennial survey which collects energy data on a sample of U.S. households. The RECS questionnaire is technical, challenging respondents' capacity to report information on energy use, equipment, and related expenses. Householders who do not pay their energy bills directly, roughly 10% of the population, are more prone to sources of response error: difficulty understanding an energy question, not having access to information, and inability to produce an accurate response. EIA fills this quality gap by conducting a supplemental rental agent survey (RAS) for these households. We examine 2009 RECS data in households with both a household and a RAS interview to show patterns of response and nonresponse bias. We discuss the impact on key estimates and energy models, the utility of this approach in terms of the survey cost-error tradeoff, and the value of further tailoring the scope of the RAS in future iterations of RECS. Since the RAS is an abbreviated version of the household survey, it offers insights to the nature of response error and the value of using a targeted, multiphase survey approach.

Eliciting Illicit Work. Item Count And Randomized Response Technique Put To The Test.

Antje Kirchner, Institute for Employment Research; ✦Ivar Krumpal, University of Leipzig, , *krumpal@sozio.uni-leipzig.de*; Mark Trappmann, Institute for Employment Research; Hagen von Hermanni, University of Leipzig

Key Words: Sensitive Questions, Item Count Technique, Randomized Response Technique, Illicit Work

We address an ongoing debate how to assess sensitive topics in telephone surveys. Examining three existing methods and implementing one new method, we developed a module to measure illicit work and tested this in two CATI studies (both conducted in 2010). In an experimental setting, we compare a double-list implementation of the Item Count Technique (ICT) with direct questioning as well as a forced-response implementation of the Randomized Response Technique (RRT) with direct questioning. In the first study (ICT; n=1.603), respondents were selected from the German general population. In the second study (RRT; n=3.211), respondents of two specific populations were sampled from a register: employed persons and those qualifying for basic income support in Germany. Goal of the studies is to evaluate which method elicits more socially undesirable answers in the context of illicit work and moonlighting, particularly with regard to the specific mode of data collection and different subpopulations. Furthermore, we developed a novel method which can be applied to the measurement of sensitive metric variables. This method requires no randomizer and can be easily administered in CATI surveys.

General-Specific Questions In Survey Research: A Confirmatory Factor Analysis Approach

Fred Lorenz, Iowa State University; Virginia Lesser, Oregon State University; ✦Laura Hildreth, Iowa State University, , *lahildre@iastate.edu*; Ulrike Genschel, Iowa State University

Key Words: survey research, part-whole questions, general-specific questions, serial correlation, structural equation modeling

Questionnaires often contain "general - specific" questions about the quality of community life in which a general question (G) either precedes (GS) or follows (SG) a series of specific items (S) related to the general question. One hypothesis is that the R2 is larger in a regression of G on the specific items in the SG conditions as respondents interpret G as a call to summarize. The magnitude of the regression coefficients can provide insight into the importance of specific items, but interpretation is difficult due to collinearity. An alternative approach is confirmatory factor models where each item is a manifestation of one or more latent variables. In this model, the same insight can be gained by comparing factor loadings in the GS and SG conditions. In addition, insight can be gained into the processes by which people answer questions. We hypothesize that respondents answer questions in an orderly sequence, represented as first order serial correlations between adjacent items. Our results support this hypothesis as evidenced by an improvement in the fit of the model over what one would expect if pairs of items were correlated at random.

580 Frames, Coverage error, and Optimal Samples ■

Section on Survey Research Methods

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

How Large Should A Statistical Sample Be?

✦Violeta Cruz Menil, Hostos Community College/CUNY, 500 grand Concourse, Rm. B408, Bronx, NY 10451 US, *menilv2@aol.com*; Ruili Ye, Hostos Community College/CUNY

Key Words: population size, precision, confidence level, sample estimate

The aim of this study was limited to determining various sample sizes (n) when estimating population proportion p. Tables were generated using a C++ program which depends on population size (N), degree of precision, or error level (E), and confidence level (Z). Nineteen different population sizes, five degrees of precision and three levels of confidence were utilized. The study found out that the larger the population size, (N), the smaller the degree of precision (E) and the higher the probability/confidence level (Z), the larger the sample size must be. Two values for the sample estimate of the population proportion, p were used in this study. Practical applications of randomly pulling appropriate number of samples from huge data sets were also discussed.

Adjustment Of Sample Size In Longitudinal Surveys

✦Andrew Vogt, Georgetown University, Department of Mathematics and Statistics, Georgetown University, Washington, DC 20057-1233, *vogta@georgetown.edu*; Dhiren Ghosh, Synectics for Management Decisions, Inc.

Key Words: time series modeling, precision, process error, optimal sample size

In a repeated sample survey of a population a statistician may develop a time series model of how a population parameter varies with time. The model typically involves autoregression, drift, process error, and similar notions. Associated with the model will be model parameters that are estimated and updated after each round of sampling. Of interest is how to choose the sample size for the next survey. If the model has been accurate, one may be able to reduce the sample size in the next round; if the model has performed poorly, an increase in sample size may be needed. Policy experts may request a certain precision in the next sample estimate. On the basis of parameter estimates to date, the statistician can tell them what sample size is necessary to achieve that precision. In some cases a smaller sample size than in the past can be used with a savings in expense, and in others the sample size and expense must increase. In general a good model will lead to convergence toward an optimal sample size for the desired precision. We consider several examples of time series models and indicate how sample size can be determined at a given epoch.

Is Housing Unit Undercoverage Random?

✦Timothy L. Kennel, US Census Bureau, DC , *Timothy.L.Kennel@census.gov*; Xijian Liu, US Census Bureau; ✦Timothy L. Kennel, US Census Bureau, DC , *Timothy.L.Kennel@census.gov*

Key Words: Address Based Sampling, Coverage, Housing Unit, Frame Creation, Master Address File, Listing

The coverage of many Address Based Sampling Frames has been well documented; however, much less attention has paid to the mechanisms generating coverage errors. A clearer understanding of the mechanisms generating undercoverage can help us more effectively target areas needing improvement or listing. To investigate correlates of Housing Unit undercoverage on Address Based Sampling Frames, we used preliminary results of the US Census Bureau's Address Canvassing to assess the coverage of several Address Based Sampling Frames. We then explored the relationship between various block characteristics to the Housing Unit undercoverage rate. Our results indicate that within a universe of blocks needing coverage improvement, Housing Unit undercoverage is mostly random. We argue that the biggest predictor of block level undercoverage is the number of housing units in the block. Thus, larger blocks within a general coverage improvement universe should be given priority when conducting coverage improvements.

Comparing Cell Phone And Internet Coverage Use In The Usa And South Korea

✦Eun-Hee Choi, Dongguk University, 26,3 Pil-dong, Jung-gu, Seoul, 100-715 South Korea, *skyanjel5240@naver.com*; Sun-Woong Kim, Dongguk University; Chan-Hyoung Cho, Korea Internet & Security Agency; Mick Couper, University of Michigan

Key Words: Internet, cell phone, landline, communication, activities

Many surveys have relied on data collection methods such as landline telephone interviews, Web surveys, and recently cell phone interviews. Much of the research on these modes of data collection comes from the USA or Western Europe. To what extent do the findings apply to other countries, especially those in Asia where both cell phone and Internet penetration is high? This study compares the USA and South Korea, both countries with high levels of cell phone and Internet penetration. We analyze data from comparable surveys on information and communication technology (ICT) adoption and use in the two countries, in particular examining whether the demographic correlates of access and use often found in the USA also hold for South Korea. We will compare the relative strength of covariates of technology use in the two countries, with the goal of identifying similarities and differences in technology adoption. We discuss the implications of these analyses for surveys conducted in the two countries using various modes of data collection, and especially in the extent to which methodological findings from one country can be generalized to others.

Rdd Unplugged: Findings From A Household Survey Using A Cell Overlap Design

John Hall, Mathematica Policy Research, Inc. ; ✦Barbara Lepidus Carlson, Mathematica Policy Research, 955 Massachusetts Ave., Suite 801, Cambridge, MA 02139, *bcarlson@mathematica-mpr.com*

Key Words: RDD, cell phone, dual-frame, sampling

Sampling the U.S. residential population using list-assisted random digit dialing (RDD) of landline telephone numbers has become problematic due to the increasing proportion of the population that is reachable only through cell phones. To address this coverage problem, round 6 of the Health Tracking Household Survey (HTHS6) employed an RDD dual-frame "cell overlap design": sample was selected from landline and cell frames, and interviews attempted with all contacted households. Other approaches sometimes used to address the coverage issue include address-based sampling and dual-frame RDD designs where the cell frame is screened for cell-only households. HTHS6 asked a series of questions about telephone usage from respondents in both the landline and cell sample frames. This paper will discuss contact and cooperation rates, and number of calls per complete, by sample frame. In addition, this paper will provide information about landline and cell telephone usage by sample type, and compare characteristics among the various telephone usage categories (cell-only, cell-mostly, landline-mostly, landline-only), including demographics, health status, and insurance coverage.

Goodness Of Fit Tests In Dual Frame Surveys

✦Yan Lu, university of new mexico, Mathematics and Statistics Department, University, of New Mexico, Albuquerque, NM 87131, *luyan@math.unm.edu*; Sharon Lohr, Arizona State University

Key Words: dual frame surveys, longitudinal quantities, goodness of fit tests, simulations

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 reviewed. Goodness of fit (GOF) tests to assess the model fit are developed. Statistical properties of the GOF tests are investigated. Simulation studies and real data example are given to illustrate the developed tests.

Using The 2009 Nhts Cell Phone Only Sample To Examine Travel Behavior

✦Adella Santos, Federal Highway Administration, Office of Policy Information, 1200 New Jersey Ave., SE, Rm E83-426, Washington, DC 20590, *adella.santos@dot.gov*; Nancy McGuckin, Travel Behavioral Associates; Robert Santos, The Urban Institute

Key Words: Cell Phone Only, Travel Behavior, National sample, Non coverage bias

Abstract For four decades the National Household Travel Survey (NHTS) has provided passenger travel trends of the American Public. The survey has been used to examine the travel behavior in issues that address congestion management, safety planning, travel modeling demands, air quality analysis, performance measures of system uses and travel trends. In the lastest series, the 2009 NHTS included a complex RDD telephone sample of 150,000 household interviews with a small Cell Phone Only (CPO) household sample of 1250. It was the first time a national CPO travel household survey was conducted in the US. The main purpose of the CPO sample was to gain knowledge on how CPO households travel differently from land line households. As the proportion of households in the U.S. steadily increases over time, it is important to explore potential noncoverage bias on key travel behavior items. This paper will compare the NHTS CPO and RDD samples to assess the possibility of non-coverage bias and it will focus on key demographics subgroups known to have higher rates of cell phone only access. Our findings will provide insight into the design of future rounds of the NHTS.

581 Social Statistics in Al Aspects of Life

Social Statistics Section, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Fear Of Crime: Analyzing The National Crime Victimization Survey

✦Elke Amelie Moons, Statistics Netherlands, Postbus 4481, Heerlen, International 6401 CZ The Netherlands, *e.moons@cbs.nl*

Key Words: crime victimization survey, fear of crime, safety

'Fear of crime' has become an often discussed topic in the criminology literature and it is frequently used by policy-making officials too. However, this 'fear of crime' has become a container concept over the years, so in order to analyze it properly it is necessary to make a distinction between its three different components: the cognitive aspect (i.e. the perceived risk), the affective aspect (i.e. feelings of unsafety) and the behavioral aspect (avoidance, prevention measures). The Dutch Crime Victimization Survey (DCVS) that is carried out on a yearly basis by Statistics Netherlands contains several questions that may help to determine 'fear of crime'. This paper focuses on 'fear of crime' concerning (attempted) burglary. The data are those of the 2009 DCVS. Several statistical techniques are applied to determine whether there exists a difference in the three concepts determining 'fear of crime' on the one hand and on the other hand between individuals that have been victims of (attempted) burglary in the last year, in the past five years or not. Differences in socio-demographical features are also accounted for.

Religious Affiliation And The Numbers Of Sexual Partners Of U.S. Men And Women

✦Hee-Choon Shin, NORC at the University of Chicago, 1155 E. 60th St., Rm. 349, Chicago, IL 60637, *shinh@uchicago.edu*

Key Words: Binomial, Poisson, Chi-Square, Religion, Sexual Behavior

According to Max Weber (1905), the main factor for the development of capitalism in Northern Europe was the Prostestant work ethic. In this paper we will examine the relationship between religious affiliation and the numbers of sexual partners of U.S. men and women. The first purpose of this paper is to describe and compare the distributions of numbers of current (during 12 months) sexual partners among U.S. women and Men by religious affiliation. After considering the univariate distributions, bivariate relationships between the number and other relevant explanatory variables (religious, demographic, and social characteristics) will be examined. Finally, the number of sexual partners among U.S. women and men is expressed as a function of religious affiliation and other explanatory variables respectively. The General Social Surveys (GSS) and the National Survey of Family Growth (NSFG) will be analyzed. The General Social Surveys (GSS) and the National Survey of Family Growth (NSFG) will be analyzed. Preliminary analyses show a higher number of sexual partners among "no religion" respondents as compared to Protestant/Catholic/Other Region respondents.

Language And Immigrant Status Effects On Disparities In Hispanic Children'S Health Status And Access To Health Care

✦Rosa Avila, National Center for Health Statistics, 3311 Toledo Rd., Room 2121, Hyattsville, MD 20782, *ravila@cdc.gov*; Matthew Bramlett, National Center for Health Statistics

Key Words: Child Health, Disparities, National Survey, Immigrant status, Primary household language, Hispanic children

Our objective is to estimate Hispanic/non-Hispanic disparities in health status and access to care and assess the extent to which disparities can be explained by household primary language and immigrant status. Using the 2007 National Survey of Children's Health, funded by the Maternal and Child Health Bureau and conducted by the National Center for Health Statistics, we calculated disparities for various health indicators between Hispanic/non-Hispanic children with logistic regression predicted marginals adjusted for socio-demographic factors, and recalculated the disparities by including household language and immigrant status in additional models. Controlling for language and immigrant status greatly reduces Hispanic/non-Hispanic health disparities, although it does not completely eliminate all disparities showing a poorer outcome for Hispanic children. Adjusted estimates for English-speaking and nonimmigrant Hispanic children are similar to those of non-Hispanic children for most health indicators. Hispanic/non-Hispanic health disparities are largely driven by that portion of Hispanic children who are either newly-arrived in the US and/or do not speak primarily English at home.

Analyzing The Duration Of The Removal Process For Illegal Immigrants

✦Anthony Kasskekert, Office of Immigration Statistics, 800 K Street NW, 10th Floor, Suite 1000, Washington, DC 20528, *anthony.kassekert@dhs.gov*

Key Words: Immigration, Enforcement, Survival modeling, Bayesian

The timely removal of illegal immigrants for violations of U.S. legal statutes is important in order to reduce administrative costs, free scarce resources, and protect the public. This study provides both a descriptive analysis of return times and an initial attempt to explain the length of the removal process by statistical methods. The dependent variable is the removal process duration from the date of apprehension until the day of removal. The covariates employed to predict removal time include geography, age, gender, nationality, month of apprehension, length of stay in the U.S., conviction of an aggravated felony or a drug charge, disposition, employment history, and immigration status (i.e. legal or illegal entry into the U.S.). Frailty survival modeling using Bayesian estimation will be used to estimate the duration of removal processing. The frailty term will be used to cluster cases based on the geographic area where they are processed. A relative risk survival model will be employed if the assumptions hold. A semi-parametric approach using a generalized additive survival model will be attempted if the relative risk model assumptions fail.

A Just World? Processes Of Social Stigma In Mental Illness

✦Daniel J. Yanosky II, Kennesaw State University, 1000 Chastain Road, MD #1601, Kennesaw, GA 30144 USA, *dyanosky@kennesaw.edu*

Key Words: Mental Health, Mental Disorders, Public Opinion, Structural Equation Modeling

Social stigma negatively impacts the lives of individuals suffering from mental illness directly and indirectly by inhibiting help-seeking behaviors. In order to better understand stigma, we hypothesized a number of respondent and situational variables to influence social stigma. A 48-item survey was administered to N = 466 university students in the Southeastern United States. Each survey 1) presented a vignette describing a character with mental illness, 2) asked respondents to rate their attitudes towards the character, and 3) measured respondents' level of possessing a just-world view. A just-world view is characterized by the belief that individuals create and therefore are responsible for their own condition. Variables included the gender, ethnicity, danger level, and controllability of the character/situation, and respondent age, gender, ethnicity, and just-world view. Two structural regression models were fit to the data. Results demonstrated that mentally-ill women were stigmatized less than men, no significant effect of just-world view on stigma was detected, and African American and Caucasian mentally-ill individuals were not stigmatized differently, among others.

Gender Gap in Postsecondary Education Trajectory in Science, Technology, Engineering, and Mathematics (STEM) Fields

✦Xiaolei Wang, American Institutes for Research, 1990 K St. NW, Suite 500, Washington, DC 20006, *xwang@air.org*

Key Words: gender studies, longitudinal data, STEM education

The study investigated the gender gaps in persistency in postsecondary education in Science, Technology, Engineering and Mathematics (STEM) fields, and how individual, family, and school characteristics are related to this educational outcome. Data for this study came from National Educational Longitudinal Survey. A nationally representative sample of eighth-graders was first surveyed in the spring of 1988, and then followed up in 1990, 1992, 1994, and 2000 (2000 is the year in which most sample members were 8 years removed from high school enrollment). Multinomial logistic regression methods were used to evaluate gender differences. The results of the study showed gender gaps in the likelihood of persisting in postsecondary education in STEM fields: men were more likely than women to persist in STEM fields. Such a gap remained statistically significant after high school achievement in math, reading, and science, as well as other individual, family, and high school characteristics variables were controlled for. Furthermore, results showed interaction effects in educational trajectory between gender and academic preparation variables and family variables.

A Meta Analysis Of Mega Projects In Transportation

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

Key Words: meta analysis, mega projects, transportation

This paper will be a meta-analysis of scientific research and policy papers that review mega-projects related to transportation and sustainable development in several countries. The objective is to develop best practices for small states like the State of Delaware. The results will be used for environmental and transportation decision making and, also, can become a framework for policy makers that can be adapted to future projects. Several case studies will be examined as part of the research.

582 Sample Size in Clinical Trials ■

Biopharmaceutical Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Optimal Sample Size And Go/No-Go Decision Based On Probability Of Success

✦Kaihong Jiang, Sanofi-aventis, BX2-400B, 200 Crossing Blvd., Bridgewater,, NJ 08807, *kaihong.jiang@sanofi-aventis.com*

Key Words: Go/No-go decision, Probability of success, Program-wise planning, Optimal sample size, Bayesian, Phase II-III

In the course of late stage drug development, a phase II/III program faces a series of sample size planning and decision problems: the simultaneous sample size determination (SSD) for phase II and III trials at the program planning stage, the phase II go/no-go decisions, and the SSD of phase III trials incorporating phase II data. Conventionally, SSD are carried out on a trial-by-trial basis and independent of one another even though they are within the same program; and go/no-go decisions are usually based on the p-values, point estimates and confidence intervals which could be ambiguous for decision making. In a combined empirical Bayesian and frequentist framework, a probability of success (POS) function is developed to allow an integrated approach to address all three SSD and go/no-go decision problems above for a phase II-III program in which both the phase II and III trials share a common, normally distributed response variable. In addition, as the POS function includes the conventional power function as a special case, it enjoys broader applications in trial planning than the power-based procedures, including the sample size determination of phase II trials.

Power And Sample Size Calculations With The Additive Hazards Model

✦Ling Chen, Washington University in St. Louis, 63110, *ling@wubios.wustl.edu*; Chengjie Xiong, Department of Biostatistics, Washington University

Key Words: right-censored data, power, sample size, additive hazards model

Existing methods on sample size calculations for right-censored data largely assume the failure times follow exponential distribution or the Cox proportional hazards model. Such methods under the additive hazards model are scarce. Motivated by a well known example of right-censored data which the additive hazards model fits better than the Cox model, we proposed a method for power and sample size calculation for a two-group comparison assuming the additive hazards model. This model allows the investigator to specify a group difference in terms of a hazard difference and choose increasing, constant or decreasing baseline hazards. The power computation is based on the Wald test. Extensive simulation studies are performed to demonstrate the performance of the proposed approach. Our simulation also shows substantially decreased power if the additive hazards models is misspecified as the Cox proportional hazards model.

R Functions for Sample Size and Power Calculations for Assessing Consistency of Treatment Effects in Multi-Regional Clinical Trials

✦Mingyu Li, Celgene, , *mli@celgene.com*; Hui Quan, Sanofi-Aventis; Joshua Chen, Merck; Yoko Tanaka, Eli Lilly; Peter Ouyang, Celgene; Xiaolong Luo, Celgene Corporation; Gang Li, J&J Pharmaceutical Research and Development, L.L.C.

Key Words: similarity, random effect, unconditional and conditional power, interaction test

Multi-regional clinical trials have been widely used for efficient global new drug developments. Due to potential heterogeneity of patient populations, it is critical to evaluate consistency of treatment effects across different regions in a multi-regional trial in order to determine the applicability of the overall treatment effect to the patients in individual regions. Quan et al. (2010) proposed definitions for the assessments of consistency of treatment effects in multi-regional trials. To facilitate the application of their ideas to design multi-regional trials, in this paper, we provide the corresponding R functions for calculating the unconditional and conditional power for demonstrating consistency in relationship with the overall/regional sample sizes and the anticipated treatment effects. Detailed step by step instructions and trial examples are also provided to illustrate the applications of these R functions.

One-Sample Proportion Testing Procedures For Hypothesis Of Inequality

Wayne Zhong, Octagon Research Solutions, Inc.; ✦Bob Zhong, Johnson and Johnson, 200 Great Valley Parkway, Malvern, PA 19355, *bzhong@its.jnj.com*

Key Words: Single arm, Phase 2 clinical trial, Hypothesis of inequality, exact binomial distribution, minimized two-stage design

The primary objective of a phase II single arm clinical trial is to determine whether a new treatment is of sufficient activity for a disease to warrant further development. This paper presents one-stage and two-stage designs for testing the response rate of a test treatment against the response rate of an existing treatment. The null and alternative hypotheses are that the response rate of the test treatment are equal and unequal to that of the existing treatment, respectively. The testing procedures are calculated with exact binomial distribution. Two-stage designs presented minimize total sample size and satisfy type I and II error constraints, with flexibility in the first stage sample size.

Sample Size Determination For Cut Point Of Immunogenicity Assays

Lanju Zhang, MedImmune; ✦Jianchun Zhang, Medimmune, LLC, One Medimmune Way, , *zhangji@medimmune.com*; Harry Yang, MedImmune

Key Words: Confidence interval, Nonparametric, Quantile

Immunogenicity assays are critical for evaluation of safety profiles of biological drugs. Key to the development and validation of an immunogenicity assay validation is the determination of a cut point. Precise estimation of this parameter requires adequate number of test samples. In practice, data for cut point calculation may have a normal, gamma, or any other skewed distributions. This talk centers on sample size determination when the cut point is the defined as a quantile of the underlying distribution.

Power Calculation In Clinical Trials

✦Lingling Han, Merck, 126 E. LINCOLN AVENUE, Rahway, NJ 07065, *lingling_han@merck.com*

Key Words: Power, Incomplete block design, clicial trials

Power calculation is very important in study designs in clinical trials. In this talk, Power calculation in different designs for different hypotheses will be discussed. Software for power calculation will be reviewed and compared. Particular, power calculation for incomplete block designs will be presented by using both formula and simulation.

583 Testing of Hypothesis: An Ensemble of Examples

ENAR

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

An Alternative Test Of Location Parameter

✦Chand K Chauhan, Indiana Purdue University, 2101 E Colesium Blvd, Fort Wayne, IN 46805, *chauhan@ipfw.edu*; Yvonne M Zubovic, Indiana Purdue University

Key Words: percentiles

In this paper the authors discuss a test for the mean of a normally distributed population when only two percentile values, such as the seventy fifth and the twenty fifth, are known. The suggested test may be useful when only two percentile values of a sample are available. The distribution of the suggested test statistic can be approximated by a t distribution. Simulation study is used to calculate the performance of the proposed test and it is compared with that of a regular t test when all the data values are available. The results are derived for several different sample sizes and for several percentile values.

Null, Not Void

✦Pamela A Shaw, National Institute of Allergy and Infectious Diseases, 6700-B Rockledge Drive, MSC-7630, Bethesda, MD 20892-7630, *shawpa@niaid.nih.gov*; Michael A Proschan, National Institute of Allergy and Infectious Diseases

Key Words: Hypothesis Tests, Nuisance Parameters, Strong and Weak Null

When performing a hypothesis test, the practitioner commonly focuses on one key parameter of interest, such as the mean. Choice of the null/alternative space impacts the choice of the test and what conclusions can be drawn from this statistical test. This choice is often given little attention. When a parameter is not of interest it is common to consider it a nuisance parameter and give it no consideration when defining the alternative hypothesis. We will illustrate with real data examples that even if a parameter is not of primary interest, it can still impact important properties of the chosen test. Mathematically, the issue can often be reduced to specifying what deviation from the strong null is included in the alternative. The type of deviation that is scientifically plausible and that is of scientific interest are both important to consider. We present several examples where specification of the alternative and the statistical test involved important distinctions of a seemingly nuisance parameter in the alternative space. Cases considered include ones where without careful thought, a less powerful test or one with misleading results could easily have been chosen.

Sign-Based Regression Quantiles

Peter F Tarassenko, Tomsk State University; Alexander Zhuravlev, Tomsk State University; ✦Sergey Tarima, Medical College of Wisconsin, 8701 Watertown Plank Rd, Wauwatosa, WI 53226, *starima@mcw.edu*

Key Words: quantile regression, nonparameteric methods, regression, sign based procedures

A sign-based (SB) approach suggests an alternative criterion for quantile regression (QR) fit. We generalize the well-known median regression SB procedure to the case of an arbitrary quantile level and compare it with the least absolute deviations (LAD) method. For linear models the SB criterion is a piecewise constant function and its minimization leads to one or several multidimensional polygons. We compare the mid-point of these polygons with the LAD estimator. Asymptotic properties of LAD and SB estimators are equivalent; however, there are finite sample differences as we show in a set of simulation scenarios. In many cases of QR parameter estimation the SB procedure demonstrates a significantly lower mean squared error (MSE) as compared to LAD based estimation. For small sample sizes (up to 10) the ratio of SB and LAD MSEs can come down to 0.5. For moderate sample sizes (near 50) the relative efficiency is typically around 0.95 in the presence of heteroscedasticity. The LAD shows lower than SB MSEs for distributions close to Laplace.

Compound P-Value Statistics for Multiple Testing Procedures

✦Joshua D. Habiger, Oklahoma State University, Department of Statistics, 301-G MSCS building, Stillwater, OK 74078-1056 USA, *jhabige@okstate.edu*; Edsel A Pena, University of South Carolina

Key Words: Multiple Testing, p-value, False Discovery Rate, Family Wise Error Rate, Sample Splitting, microarray

Many multiple testing procedures make use of the p-values from the individual pairs of hypothesis tests, and are valid if the p-value statistics are independent and uniformly distributed under the null hypotheses. However, works such as Sun and Cai(2007) and Storey(2007) have shown that these types of multiple testing procedures are inefficient since such p-values do not depend upon all of the available data. This talk will provide tools for constructing compound p-value statistics, which are those that depend upon all of the available data, but still satisfy the independence and uniformity conditions. As an example, a class of compound p-value statistics for testing for location shifts is developed. It will be shown that multiple testing procedures are more powerful when applied to these compound p-values rather than the usual p-values, and at the same time still guarantee control of the desired type I error rate. Methods are used to analyze a microarray data set.

Confidence Intervals For Proportion Difference Based On Paired And Unpaired Binary Data

✦Yue Shentu, Merck & Co., 126 E. LINCOLN AVENUE PO BOX 2000, RY34-A304, Rahway, NJ 07065, *yue_shentu@merck.com*; Yabing Mai, Merck & Co.; Jiajun Liu, Merck & Co.

Key Words: paired binary data, confidence interval, missing data, proportion difference, method of variance estimates recovery

Interval estimation for the proportion difference based on paired binary data has many applications in clinical studies. Missing data in these situations often lead to incomplete paired binary data for a subgroup of subjects. We propose an extension of the method of variance of estimates recovery (MOVER) to construct confidence intervals (CIs) for the correlated proportion difference based on paired and unpaired data. Two sets of CI estimators, one based on paired data, the other based on pooled paired and unpaired data, are used to construct the asymptotic CI. Extensive simulations show that the extended MOVER-type estimator performs well under various degrees of missingness and correlations, even with small to moderate sample sizes. A real example is used to demonstrate the proposed method.

Empirical Likelihood Based Change Point Detection Approach

✦Yijie Xue, Department of Statistics, University of Georgia, 101 Cedar Street, Athens, GA 30602, *yijie.xue@gmail.com*; Nicole Lazar, Department of Statistics

Key Words: Empirical likelihood, Change point detection

Change point detection problems are often encountered in quality control and climatological data. The proposed empirical likelihood based approach provides an alternative method to solve this problem. Based on the empirical likelihood distribution with the weights estimated from empirical likelihood ratio, we set a threshold to identify potential change points. Because the empirical likelihood distribution with the estimated weights is asymptotically unbiased for the true distribution, researchers do not need to specify the model of the data, that is, it is model-free. The proposed method will be compared to existing methods via simulations.

Goodness-Of-Fit Methods Of Finite Mixture Models Based On Cumulative Pseudo-Residuals

✦Junwu Shen, Merck & Co., Inc., 2015 Galloping Hill Road, K-15-2-2445, Kenilworth, NJ 07033, *junwu.shen@spcorp.com*; Shou-En Lu, University of Medicine and Dentistry of New Jersey

Key Words: Goodness-of-fit, Mixture models, Cumulative residuals

One of the challenges in the application of finite mixture models is to evaluate the adequacy of the model. Many different methods including likelihood ratio test, AIC, and BIC have been used extensively to compare the relative goodness-of-fit for different mixture models. However, there was limited literature on how to evaluate the discrepancy between observed data and a specific mixture model. In this paper, we proposed a model checking technique extended from the principle of cumulative residuals (D.Y. Lin, et al. 2002) to evaluate the goodness-of-fit for a mixture regression model. Using this method, both the component means and the mixing proportions modeled by linear and logistic regressions can be tested separately and jointly. For each of these tests, we can both visually and numerically evaluate the functional form of a covariate or the link functions. Simulation studies showed that the proposed tests perform well in terms of type I error and have a reasonable power. In addition, the proposed goodness-of-fit tests can also be extended to mixture models with random effects. The proposed method was applied to the data analysis for an environmental health study.

584 Analysis of Binary and Binomial Data ■

Biopharmaceutical Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Sample Size Determination For Alternate Periods Of Use Study Designs With Binary Responses

✦Jorge Morel, The Procter and Gamble Company, 6280 Center Hill Avenue, Cincinnati, OH 45224, *morel.jg@pg.com*; Nagaraj K. Neerchal, Department of Mathematics and Statistics

Key Words: Cross-over Designs, Repeated Measures, Menstrual Studies, Diaper Studies, Exchangeable Correlation, Power

We consider several study designs that arise in practice, which are variations of standard cross-over designs. Often, they may result from modifications of standard cross-over designs due to practical considerations. Characteristic features of these studies are: a) treatments consist of external use of products with little or no possibility of carry over effects, and b) the periods of use are dictated by the subjects or by some specific event such as diaper leakage or menstrual flow. We consider a number of such study designs for estimating the difference in the efficacy of two treatments or test products. We provide brief descriptions of studies to motivate the study design, the underlying data structure, and computations of the variances of the usual unbiased estimators of the difference in efficacy, and the sample size formulas. The situations considered here cover a number of popular cross-over designs. Objective of our work is to provide guidance to a wide audience on how to answer the sample size question for their own non-standard situations. We conclude with a simulation study conducted to investigate the impact of estimation on the sample size determination.

Empirical Study Of A Measure Of Reduction For Two Binomial Variates

✦Serena Petersen, University of Nevada, Las Vegas, 4505 Maryland Parkway Box 454020, Las Vegas, NV 89154-4020, *cho@unlv.nevada.edu*

Key Words: Measure of reduction, Maximul likelihood estimation, Asymptotic unbiasedness, Asymptotic variance, Confidence interval

We study a measure of reduction (MOR) for two independent binomial variates. An estimator for the MOR, denoted by rho(?), is considered from a modified maximum likelihood estimator. We investigate the asymptotic behavior of its unbiasedness and the variance of the estimator. The normality of the is established and approximately normally distributed. Based on the normality of the measure rho (?), we constructed a confidence interval. Monte Carlo simulation is carried out for the various scenarios and to examine its finite sample behavior.

Bounds On The Proportion Of Negative Response In Cross-Over Designs With Binary Outcomes

✦Edwin Andong NDUM, ACT, Inc., 1355 Ozark Ridge, Coralville, IA 52241, *ean_dong@yahoo.co.uk*; Gary L. Gadbury, Kansas State University

Key Words: Bounds, Crossover design, Potential outcomes, Probability of negative effect, Causal effect, binary outcomes

Cross-over designs allow for a subject's outcome on more than one treatment to be observed over multiple periods. However, it is impossible to measure a true individual causal effect from observed data because only one outcome response to a treatment is observable at any particular time. Often, the difference between the proportion of subjects benefiting on a new and standard treatment is interpreted as the probability that an individual will benefit on the new treatment. This interpretation can be misleading when the effects of a treatment vary widely across subjects, i.e., subject-treatment interaction is present in the population. This talk considers bounds for a nonestimable probability of negative individual effect when the average effect is positive. The potential outcomes framework is used to quantify consequences of subject-treatment interaction. Initially, we construct bounds from the crossover design, TC-CT, with binary outcomes. Subsequently, we find tighter bounds using the TC-CT-TT-CC design in which some subjects are allowed to stay on the same treatment across the two periods. Concepts are illustrated with an example.

Evaluating A Difference In Proportions In The Presence Of Heterogeneity: An Application To Integrated Summaries Of Safety

✦David Radley, Merck Research Labs, PO Box 1000, North Wales, PA 19454, *david_radley@merck.com*; Michael Dallas, Merck Research Labs; Adeniyi Adewale, Merck Research Labs

Key Words: Heterogeneity, stratification, binary data, integrated summaries, safety data

In integrated summaries of safety (ISoS), there is often interest in the difference between a test group and a control group in the proportion of subjects reporting a safety outcome (occurrence of a specific adverse experience, for example). When combining evidence across multiple studies data are often simply pooled; however, heterogeneity may be introduced according to study characteristics. We simulated stratified binary data, in the presence of heterogeneity of varying degrees, and obtained the point and interval estimates of the overall risk difference using several methods. The methods considered include Mietinnen & Nurminen (MN) method for stratified data using various weighting schemes; a logistic mixed model with random effects for study; and a quasi-likelihood approach. Depending on the degree of heterogeneity present, we provide recommendations for the best method.

Confidence Intervals For Binomial Proportions With Applications To Product Safety Signaling In Drug Exposure Registries

✦Paul Schuette, FDA CDER, 10903 New Hampshire Avenue, WO 21 Room 4659, Silver Spring, MD 20903, *paul.schuette@fda.hhs.gov*; Fiona Callaghan, NIH; C. George Rochester, FDA CDER

Key Words: binomial proportion, confidence interval, pregnancy registries

Coverage behavior for confidence intervals for binomial proportions have been investigated by various authors including Vollset (1993), Newcomb (1998), Agresti et al (1998, 2001), and Brown, Cai and DasGupta (2001). A general consensus is that the commonly used Wald confidence interval is flawed, particularly if the proportion is near 0 or 1, but agreement on a replacement is lacking. Previous investigations in this area have focused on the aggregate behavior of the confidence interval coverage across all possible values of the binomial proportion. Most recommended methods either have under coverage issues if the proportion is near 0 or 1, or else are excessively conservative. We shall compare a variety of confidence intervals based on asymptotic, exact, and quasi-exact methods when the rate of success is near 0, which is the usual context of adverse event signaling in drug exposure registries. Upon examination of the coverage probabilities for a range of sample sizes n and p near 0, we find that recommended methods such as Wilson's method have notable deficiencies, and present alternatives.

Performance Of Corrected Profile Likelihood Confidence Interval For Binomial Paired Incomplete Data

✦Sandeep M Menon, Pfizer Inc/ Boston University, 73 Walnut Street, Unit Number 5, Newton, MA 02460 USA, *sandeep_menon9@yahoo.com*

Key Words: confidence interval, binomial, paird, incomplete, coverage probability

Clinical trials often use paired binomial data as their clinical endpoint. The confidence interval is frequently used to estimate the treatment performance. Tang et al (2009) have proposed exact and approximate unconditional methods for constructing confidence interval in the presence of incomplete paired binary data. The approach proposed by Tang et al could be overly conservative with large expected confidence interval width (EXIW) when the sample size are large which is true in most clinical trials. We propose a profile likelihood method with a Jeffrey's prior correction to construct the confidence interval. This approach generates confidence interval with much better coverage probability and shorter expected confidence interval widths. SAS codes to execute this method is also presented.

585 Distribution Theories

IMS

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Exact Sampling For Intractable Probability Distributions Via A Bernoulli Factory

✦James M. Flegal, University of California, Riverside, 1428 Olmsted Hall, 900 University Ave, Riverside, CA 92521 United States, *jflegal@ucr.edu*

Key Words: Markov chain, Monte Carlo, Perfect sampling, Gibbs sampler

Many applications in the field of statistics require Markov chain Monte Carlo methods. Determining appropriate starting values and run lengths can be both analytically and empirically challenging. A desire to overcome these problems has led to the development of exact, or perfect, sampling algorithms which convert a Markov chain into an algorithm that produces i.i.d. samples from the stationary distribution. Unfortunately, very few of these algorithms have been developed for the intractable distributions that arise in statistical applications, which typically have uncountable support. Here we study an exact sampling algorithm using a geometrically ergodic Markov chain on a general state space. Our work provides a practical implementation of a previously studied rejection sampling approach. To this end, we provide an explicit bound for the proposal distribution and implement the Bernoulli factory. We illustrate the algorithm on a bivariate Gibbs sampler.

Triangular-Weibull Distribution And Its Relation To The Other Generalizations

✦Gokarna R. Aryal, Purdue University Calumet, 2200 169th Street, Hammond, IN 46323, *aryalg@purduecal.edu*

Key Words: Weibull Distribution, Triangular Distribution, Probability Distribution

The Weibull distribution has been extensively applied to lifetime data analysis. Various generalizations of the Weibull distributions has been developed and applied in several areas. In this study we present a generalization of the Weibull distribution through an application of the probability integral transform using triangular ~[0,1] distribution. A comparison with other generalizations of the Weibull distribution will also be presented.

Zero Inflated Negative Multinomial Distribution

✦Santanu Chakraborty, University of Texas - Pan American, Department of Mathematics, 1201 West University Drive, Edinburg, TX 78539, *schakraborty@utpa.edu*

Key Words: Negative Multinomial, Zero Inflated, Failure Vector

Negative Multinomial Distributions do exist in the literature for about half a decade now. This is considered if there are possibilities of more than one type of failures before achieving the required number of successes. These different types may be put together in the form of a vector called the failure vector where each vector component indicates a particluar type of failure. Now it may so happen that many of these types of failures occur very rarely before getting the desired number of successes in real life situations. So, some of the components in the failure vector could be zeros. In such situations, it makes sense to talk about Zero Inflated Negative Multinomial (ZINM) distributions. Zero Inflated Negative Binomial(ZINB) distributions already exist in the literature where there is only one type of failure and the number of failures is zero more often than a usual negative binomial distribution. There have been a lot of studies already as regards fitting Zero Inflated Negative Binomial models to count data sets. In this work, we formally introduce ZINM and also touch upon some Frequentist and Bayesian inferential issues for this distribution.

Characterization Of Binomial Distribution And Application To Goodness-Of-Fit Test

✦Dhanuja Kasturiratna, Northern Kentucky University, Dept. of Mathematics and Statistics, Highland Heights, KY 41099, *kasturirad1@nku.edu*; Truc T Nguyen, Bowling Green State University; Arjun K Gupta, Bowling Green State University

Key Words: Characterization, Goodness of fit test, Binomial distribution

Characterization of binomial distribution based on second conditional moment has been obtained. This characterization is used to construct a conditional exact goodness-of-fit test for testing the binomial distribution. The approximation of this test in the large sample case is shown to be a two sided chi square test for binomial distribution. The powers of the tests are studied by Monte Carlo methods for several alternatives.

Introducing A Generalized Bivariate Distribution For Count Data

✦Kimberly F. Sellers, Georgetown University, 306 St. Mary's Hall, Department of Mathematics and Statistics, Washington, DC 20057, *kfs7@georgetown.edu*; Narayanaswamy Balakrishnan, McMaster University

Key Words: bivariate Poisson distribution, bivariate geometric distribution, bivariate Bernoulli distribution, Conway-Maxwell-Poisson, dispersion

The bivariate Poisson distribution is a popular distribution for modeling bivariate count data. Its associated assumptions, however, may prove limiting in some contexts. To allow for data dispersion, we develop a bivariate Conway-Maxwell-Poisson (COM-Poisson) distribution that includes the bivariate Poisson, bivariate geometric, and bivariate Bernoulli distributions as special cases. As a result, the bivariate COM-Poisson distribution serves as a flexible alternative for modeling bivariate count data, particularly in the presence of data dispersion.

On Semiparametric Modelling Of Heavy Tailed Distributions

✦Dieter Schell, University of Konstanz, Markgrafenstrafle 67, Konstanz, 78464 DE, *dieter.schell@uni-konstanz.de*; Jan Beran, University of Konstanz

Key Words: heavy tails, tail index

A combination of a Pareto density and suitable basis functions is used to define a large class of heavy-tailed distributions. The corresponding coefficients are estimated simultaneously with the tail index. Consistency and asymptotic normality are derived. Moreover, estimation based on an increasing number of basis functions is considered. The resulting class of tail index estimators is compared to optimal nonparametric methods known in the literature. Simulations and data examples illustrate the theoretical results.

Ensemble Minimax Estimation For Multivariate Normal Means

Lawrence Brown, University of Pennsylvania; ✦Hui Nie, University of Pennsylvania, 3730 Walnut Street, Suite 400, Philadelphia, PA 19104, *niehui@wharton.upenn.edu*; Xianchao Xie, Harvard University

Key Words: Shrinkage, Empirical Bayes, Ensemble Minimax, James-Stein Estimation, Random Effects Models

This article discusses estimation of a heteroscedastic multivariate normal mean in terms of the ensemble risk. We first derive the ensemble minimaxity properties of various estimators that shrink towards zero. We then generalize our results to the case where the variances are given as a common unknown but estimable chi-squared random variable scaled by different known factors. We further provide a class of ensemble minimax estimators that shrink towards the common mean.

586 Regression Methods

Business and Economic Statistics Section

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Least Absolute Deviation Estimation for General Autoregressive Moving Average Time Series Models

✦Rongning Wu, Baruch College, One Bernard Baruch Way, Box 11-220, New York, NY 10010, *rongning.wu@baruch.cuny.edu*; Richard A Davis, Columbia University

Key Words: autoregressive moving average model, least absolute deviation estimation, noncausality, noninvertibility

In this paper, we study least absolute deviation (LAD) estimation for general autoregressive moving average (ARMA) time series models that may be noncausal, noninvertible or both when the underlying noise is non-Gaussian. For stationary ARMA models with Gaussian noise, causality and invertibility are assumed in order for the parameterization to be identifiable. The assumptions, however, are not required for models with non-Gaussian noise, and hence are removed in our study. We deconstruct a stationary ARMA model into its causal, purely noncausal, invertible and purely noninvertible components, and formulate the proper LAD objective function. Following the approach taken by Davis and Dunsmuir [Least absolute deviation estimation for regression with ARMA errors, Journal of Theoretical Probability 10 (1997) 481-497], we derive a functional limit theorem for random processes based on the LAD objective function, and establish the consistency and asymptotic normality of the LAD estimator. The performance of the estimator is evaluated via simulation and compared with the asymptotic theory. Application to a financial time series consisting of volumes of traded stocks is also provided.

Median Regression For Sur Models With The Same Explanatory Variables In Each Equation

✦Zangin A.A. Zeebari, Jonkoping International Business School, Gjuterigatan 5, JIBS, P.O Box 1026, Jonkoping, 55111 Sweden, *zangin.zeebari@jibs.hj.se*; Ghazi Shukur, Jonkoping International Business School

Key Words: Median SURE regression, robustness, efficiency, SURE models

In this paper we introduce an interesting feature of the Generalized Least Absolute Deviation (GLAD) method, that contrary to the collapse of GLS estimation of Seemingly Unrelated Regression Equations (SURE) to the OLS estimation when the same regressors are common between equations, the estimation of the proposed methodology does not collapse to the equation-by-equation LAD estimations. This is important since contrary to the least squares method, one can take advantage of efficiency gain due to cross-equation correlations even if the system includes the same regressors in each equation. Given the existence of systemwise medians of the errors from different equations when the cross-equation errors are correlated and regressors are the same for each equation, we show by a Monte Carlo experiment, an empirical study and analytical results that the proposed method does not collapse to the LAD estimation and it is more robust and more efficient than LAD estimation of the separate median regression equations and GLS method as well.

Fitting Mixtures Of Logit Regressions With The Forward Search

✦Margherita Velucchi, Department of Statistics - University of Florence _, V.le G.B. Morgagni,59, Firenze, International 50133 Italy, *velucchi@ds.unifi.it*; Matilde Bini, Department of Economics - European University of Rome -

Key Words: Forward search, Robustness, Logit Regression, Firms Behavior

The forward search is a powerful general method to detect multiple outliers and to determine their effect on inferences about models fitted to data. From the monitoring of a series of statistics based on subsets of data of increasing size we obtain several views of any hidden structure. One of the problems of the forward search has always been the lack of an automatic link among the great variety of plots which are monitored. Sometimes, some features emerge unexpectedly during the progression of the forward search only when a specific combination of forward plots is inspected at the same time. These features have to be harmonized and linked together in order to give an exhaustive description of a very complex problem. In this paper, we aim at filling this gap and use a set of new robust graphical tools whose power will be demonstrated on a logit regression problem. We use real microeconomic data (Italian firms balance sheets in 2007) and simulated data and we show the dynamic interaction with different "robust plots" to highlight the presence of groups of outliers and regression mixtures and appraise the effect that these hidden groups exert on the fitted model.

Mm Versus Ml Estimates Of Structural Equation Models With Interaction Terms: Robustness To Non-Normality Of The Consistency Property

✦Albert Satorra, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, Barcelona, 08005 Spain, *albert.satorra@upf.edu*; Ab Mooijaart, Leiden University

Key Words: maximum-likelihood, method of moments, third-order moment, asymptotic bias, latent-variable, structural relation

A standard assumption in structural equation models with interaction terms is the normality of all the random constituents of the model. In applications, however, there may be predictors, disturbance terms of equations, or measurement errors, deviating from normality. The present paper investigates how deviation from normality affects the consistency of two alternative estimators; namely, the maximum likelihood (ML) and the method of moments (MM) estimators. The ML approach requires full specification of the distribution of observable variables while this is not required in the MM approach. It will be seen that while the MM estimator is insensitive to departures from normality of all the random constituents of the model not involved in the interaction terms, such deviation from normality distorts considerably the consistency property of the ML estimator. The paper provides analytical results showing the consistency of MM when using a proper selection of moments up to order three, and presents a Monte Carlo illustration showing how the consistency of the ML estimator breaks down when there is deviation from normality. It is concluded that for a va

Partially Adaptive And Semiparametric Estimation Of Regression Models For Grouped Data

✦Jason Blaine Cook, Brigham Young University Economics Department, 445 N 400 E Apt. 329, Provo, UT 84606, *jascook5@gmail.com*; James B. McDonald, Brigham Young University

Key Words: Partially adaptive estimation, Semiparametric estimation, Grouped data, Interval data

Several valuable data sources, including the census and NLSY, include data based on interval responses. Many empirical studies attempt estimation by assuming the data correspond to the interval midpoint and run OLS regressions or performing maximum likelihood estimation (MLE). Stata provides a MLE approach, based on the assumption of normality, which allows for intra-group variation. In the presence of misspecification, these estimates are inconsistent. Other solutions include utilizing nonparametric estimation techniques. In this paper we take an intermediate position, using partially adaptive estimation, which builds on a MLE framework using flexible parametric probability density functions (pdfs). Semiparametric procedures are also considered. Applications of these methods are used to estimate determinants associated with household income based on US Census data. Monte Carlo Simulations are performed to evaluate the relative efficiency of the different methods of estimation.

Interpreting Dummy Variables in Semilogarithmic Regression Models: Exact Distributional Results

✦David E. Giles, University of Victoria, Department of Economics, P.O. Box 1700, STN CSC, Victoria, BC V8W 2Y2 Canada, *dgiles@uvic.ca*

Key Words: Semilogarithmic regression, Dummy variable, Confidence interval, Percentage change

Care must be taken when interpreting the coefficients of dummy variables in semilogarithmic regression models. Existing results in the literature provide the best unbiased estimator of the percentage change in the dependent variable, implied by the coefficient of a dummy variable, and of the variance of this estimator. We extend these results by establishing the exact sampling distribution of an unbiased estimator of the implied percentage change. This distribution is non-normal, and is positively skewed in small samples. We discuss the construction of bootstrap confidence intervals for the implied percentage change, and illustrate our results with an application involving an hedonic price index for computer disk drives.

Nonlinear Regression With Conditionally Stable Innovations: A New Definition Of Financial Contagion

✦Eric Grant Stradley, Rhodes College, 2000 N. Parkway, Memphis, TN 38112, *eric.g.stradley@gmail*; Jeff Hamrick, Rhodes College

Key Words: contagion, crisis, heavy tails, stable distribution, nonlinear regression, skewness

We develop a new notion of financial contagion, or the spread of negative characteristics from one market to another, by fitting a conditionally stable model to residuals extracted from a nonlinear regression. More specifically, we estimate the return on a dependent market given the return on an independent market using a spline-based local mean function. Then, instead of assuming that the residuals have a Gaussian distribution, we assume that the residuals are independent stable random variables when conditioned on the covariate market return. In general, the stable distribution depends on four parameters, two of which control skewness and tail heaviness. With our approach, these parameters become functions that are nonparametrically estimated. For various dependent markets, we study the change in the skewness and heaviness functions from the median to the tail of an associated covariate market return distribution. After bootstrapping the results of our function estimation, we test whether, given a value in the tail of the covariate market return distribution, the residuals are more likely to be left-skewed or heavy at the left tail than at the median of the distribution.

587 Clustering and Data Imputation

Section on Statistical Learning and Data Mining

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Enhancement Of Clustering Results With Variable Selection And Resampling Methods

✦Wenzhu Bi, Department of Biostatistics, University of Pittsburgh, 326 Parran Hall, 130 DeSoto Street, Pittsburgh, PA 15261, *web10@pitt.edu*; George C. Tseng, Department of Biostatistics, University of Pittsburgh; Julie C. Price, Department of Radiology, University of Pittsburgh; Lisa A. Weissfeld, Department of Biostatistics, University of Pittsburgh

Key Words: clustering, variable selection, resampling, imaging, feature selection, PET

Clustering can be used to identify biologically distinct subgroups from an n◊p dataset without knowledge of the true group membership. Since some variables are irrelevant to clustering and may only introduce noise, variable selection methods have recently been developed to exclude these variables and to yield more reliable and parsimonious clustering results. Recently in 2010, Witten and Tibshirani introduced a general framework for variable selection by applying a Lasso-type penalty and an L2 condition. We propose to combine Witten's method with resampling techniques, such as bootstrapping or leave-one-out resampling. The goal is to alleviate the effects of noise and outliers on the variable selection and clustering results and also to generate confidence intervals for the clustering results. The performance of the proposed method is demonstrated by simulation. We then apply the method to neuroimaging data. The focus is on the analysis of voxel-level data and the identification of a subset of voxels that can be used to classify subjects into groups. We present a PET imaging example using an unspecified radiotracer to identify groups of subjects with varying amounts of tracer.

Mixture Model Cluster Analysis For Medical Diagnosis Under Different Covariance Structures Using Information Complexity

✦Bahar Erar, University of Tennessee, 344 Stokely Management Center, 916 Volunteer Blvd., Knoxville, TN 37996 USA, *berar@utk.edu*; Hamparsum Bozdogan, University of Tennessee

Key Words: model-based clustering, mixture models, information complexity, X-ray mammographic images, information criteria

In this paper we develop and present a novel mixture-model cluster analysis technique for medical diagnosis under different covariance structures of the component densities to capture the compactness, orientation, shape, and the volume of component clusters in one expert system to handle both Gaussian and non-Gaussian high dimensional heterogeneous medical data sets to achieve flexibility in currently practiced cluster analysis techniques. We develop and score several forms of the ICOMP criterion of Bozdogan (1994, 2004) as our fitness function; (i) to choose the number of component clusters, (ii) to choose the correct component covariance matrix structure among nine candidate covariance structures, (iii) to select the optimal parameters and the best fitting mixture-model, and (iv) to determine which variables best discriminate between the clusters chosen. We demonstrate our approach on a real large data set, focusing on early detection of breast cancer using X-ray mammographic images using ranklet transforms. We show that our approach improves the probability of misclassification error over the existing methods by providing an online unsupervised statistical diagnostics toolbox.

Recursive Partitioning With A Latent Response: Application To Hiv Viral Genetic Data

✦Shannon Stock, Department of Biostatistics, HSPH, 655 Huntington Avenue, SPH2, 4th Floor, Boston, MA 02115 USA, *sheezy.stock@gmail.com*; Victor DeGruttola, Department of Biostatistics, HSPH

Clinical studies evaluating the efficacy of antiretroviral (ARV) therapy in suppressing Human immunodeficiency virus (HIV) replication frequently use viral rebound to clinically indicate treatment failure. However, since HIV RNA measurements are obtained with measurement error, there is uncertainty whether an observed increase in viral load is due to measurement error rather than viral rebound. Therefore, analyses of such studies must accommodate a latent response. In this paper we present a recursive partitioning method for a latent event that uses weights estimated via maximum likelihood to estimate the probability an event occurred between successive (in time) pairs of observations, given an observed change in outcome measurements. The only parametric assumptions our method require relate to the estimation of the outcome measurements' measurement error distribution. We evaluate the performance of our method using simulation studies, and provide an application based on HIV viral genetic data using the AIDS Clinical Trials Group (ACTG) 398 clinical study.

Data Mining Categorical Predictors With Missing Values

✦Hua Fang, University of Massachucetts Medical School, , *hua.fang@umassmed.edu*; Honggang Wang, University of Massachusetts Dartmouth

Key Words: data mining, categorical predictor, missing values, high-dimensional, multiple-imputation, simulation

Predictors with categories of no clear boundaries are prevalent in all kinds of studies, for example, low- or high- cancer/behavioral risk people, better or worse consulting procedure, contrarian or non-contrarian investors, efficient or inefficient market, etc.. A number of attributes can be used to better describe this type of categorical predictors, accordingly reduce their measurement errors and increase their predictive power in statistical hypothesis testing. To accommodate missing values in the attributes, we proposed multiple-imputation based data mining techniques to characterize categorical predictors. Theoretical illustration, simulation and cases studies will be included to demonstrate the utility of our proposed method.

Consistent Selection Of The Number Of Clusters Via Clustering Stability

✦Junhui Wang, University of Illinois at Chicago, , *junhui@uic.edu*

Key Words: Cluster analysis, Cross validation, k-means, Selection consistency, Spectral clustering, Stability

In cluster analysis, one of the major challenges is to estimate the number of clusters. Most existing approaches focus on the within- and between cluster dissimilarities measured by various forms of distance. In this talk, I will present a novel selection criterion that is applicable to all kinds of clustering algorithms, distance-based or non-distance-based. The key idea is to select the number of clusters such that the corresponding clustering algorithm has the smallest instability. The clustering instability measures the robustness of any given clustering algorithm against the randomness in the sampling. Numerical examples will be provided to demonstrate the effectiveness of the proposed selection criterion. Asymptotic selection consistency of the proposed selection criterion will also be discussed.

Dynamic Clustering Methods For Analyzing Complex Systems

✦Shawn Mankad, University of Michigan, Department of Statistics, 439 West Hall, 1085 South University Ave., Ann Arbor, MI 48109-1107, *smankad@umich.edu*; George Michailidis, University of Michigan; Andrei Kirilenko, Commodity Futures Trading Commission

Key Words: clustering, spatiotemporal, signal extraction, online algorithm, complex system, finance

Driven by advances in computer technology, businesses and organizations are increasingly able to collect data streams that capture the behavior of components in complex systems. Suppose we observe noisy snapshots, at the component level, of a complex system over time. Specifically, we observe the same features for each component at each time point. A challenging issue is to identify and interpret significant component behaviors and interactions. Further, in many cases it is impractical to store and organize the full data for retrieval and analysis. Thus, methods that do not require the full data and operate given the input of information up to a point in time are essential to the analysis of large systems. Such systems are encountered in biology, economics, transportation, among others. To gain insight into these interactions, we develop a dynamic biclustering method based on the plaid model. We introduce a regularization framework that smooths the model parameters over time, allowing us to identify persistent groups (components) and the critical features that separate them consistently over time. We illustrate this methodology with an application on electronic trading market data.

Network Community Definition And Detection

✦Shuqin Zhang, Yale University, New Haven, CT 06511, *shuqin.zhang@yale.edu*; Hongyu Zhao, Yale University

Key Words: communitiy structure, random graph

Community structure or module structure is an important issue in many different kinds of networks such as social networks and biological networks. Several types of algorithms for identifying the community structure have been proposed which include the clustering techniques, modularity optimization, spectral partitioning, k-clique percolation and some other methods. Although all these methods have been shown to have their advantages, there is no consistent definition of the community structure. Very few theoretical results for the detection methods have been given. In this paper, we give a new definition for the network community based on the random graphs. An algorithm for detecting the community structure is also proposed. Numerical experiments based on simulations and real data are given to demonstrate the efficiencies of the algorithm. Theoretical analyses of the definition and the algorithm are also given to explain and support the numerical results.

588 High-dimensional data analysis ■

Section on Statistics and Marketing

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Testing a Subset of Regression Coefficients in High-Dimensional Data Analysis

✦Hansheng Wang, Guanghua School of Management, Peking University, Beijing, International 100871 P. R. China, *hansheng@gsm.pku.edu.cn*; Wei Lan, Guanghua School of Management, Peking University ; Chih-Ling Tsai, Graduate School of Management, University of California--Davis

Key Words: High Dimensional Data, Subset Testing, Regression Analysis, Partial Correlation

Modern scientific applications always encounter ultra high dimensional data, for which the method of regression analysis has been found very useful. Such type of problem has attracted considerable attention from the past literature.The past research efforts have been focusing on parameter estimation and variable selection. Nevertheless, much less has been done for hypotheses testing under an ultra high dimensional setup. To solve the problem, we systematically investigate ultra high dimensional testing problems under a regression setup. Theoretical analysis reveals that our method is valid as long as both the sample size and the predictor dimension goes to infinity. Monte Carlo studies and numerical examples are presented to illustrate the performance of the proposed test.

Using Bayesian Approach To Model Collinear Data In Marketing Analysis

✦Min Lu, The University of Memphis, 1372 Island Town Dr., Memphis, TN 38103, *lu_min27@yahoo.com*

Key Words: Collinearity, Bayesian, marketing analysis, informative prior

Collinearity in multiple regression often inflates the standard errors of coefficient estimates. In this situation, small changes in the model or the data lead to dramatic changes in coefficient estimates. Although the predictive power of the entire model wouldn't be affected, the interpretation of individual regressor is often misleading. Marketing analysis usually involves multiple correlated factors and the need of understanding individual market factor makes solving collinearity issue high priority. The common solution includes increasing sample size, using some variable reduction technique or biased regression. This presentation will discuss the bayesian approach with informative prior to be the alternative when none of the above approach is suitable and its application in marketing research.

Applying New Developments In Dimension Reduction Techniques In Statistics To Data-Rich Marketing Research

Neeraj Bharadwaj, Temple University; ✦Yuexiao Dong, Temple University, 1810 N 13th St, 326 Speakman Hall, Philadelphia, PA 19122, *ydong@temple.edu*

Key Words: Factor analysis, Parallel analysis, Principal component analysis, Sliced inverse regression, Sufficient dimension reduction

In the presence of large amounts of data, marketing researchers have relied upon dimension reduction methods (i.e., principal component analysis (PCA) and factor analysis) to reduce the number of variables effectively. Such traditional approaches extract a few significant factors that are linear combinations of the original predictor variables, but ignore the role of the response variables. To incorporate the information from both the predictor and the response variables, Naik et al. (2000) introduced sliced inverse regression (SIR), a sufficient dimension reduction technique, into the marketing research literature. SIR provides a viable alternative to PCA type dimension reduction methods as demonstrated by Monte Carlo simulations and empirical studies. Over the last decade, there have been significant advances in sufficient dimension reduction in statistics, yet few attempts to utilize them in marketing research. In this interdisciplinary research, we seek to address the call made by Naik at al. (2000) to marketing scholars and practitioners to "improve decision making by better utilizing vast amounts of marketing information" by applying recent advances in SIR from statistics.

Tree-Based Time Varying-Coefficient Regression For Product Demand Prediction

✦Jay Wang, Hewlett-Packard Labs, 1501 Page Mill Rd, Palo Alto, CA 94304, *qqwjq9916@gmail.com*; Kay-Yut Chen, Hewlett-Packard Labs; Guillermo Gallego, Columbia University

Key Words: elasticity, splines, tree-based methods, varying-coefficient regression

Estimating the aggregated market demand for a product in a dynamic market is intrinsically important to manufacturers and retailers. Motivated by the need for a statistical demand prediction model to facilitate laptop pricing at Hewlett-Packard, we have developed a novel “partitioned regression” model that is easily interpretable and have good prediction power. The proposed method is generally applicable to situations where predictions are made based on a large number of categorical variables and a much smaller number of continuous variables. Similar to regression trees, the partitioned regression uses a “greedy algorithm” to find optimal partitions of the products, where “optimality” refers to the maximal reduction of $-2\log$ likelihood. Within each partition (or “market segment”), a time varying-coefficient regression model is built to incorporate the temporal dynamics of the demand-price relationship, to reflect the change of market volume and customer preference at different phases of product life cycle. The proposed methodology is applied to a real dataset containing laptop sales of all brands covering September 2009 to June 2010 in all five states of Australia.

Order Detection For Factor Analysis Model Using Adaptive Group Lasso

✦Liuxia Wang, Sentrana Inc., 1725 I St. NW, Ste 900, Washington, DC 20006, *liuxia.wang@sentrana.com*; Yulin Li, K&L Consulting Services Inc

Key Words: Adaptive group lasso, selection consistency, factor analysis model, ECM algorithm

One crucial issue in factor analysis (FA) model is to detect the number of factors to summarize the original data. A wealth of approaches have been proposed for this matter. In this paper, we formulate the order detection for FA model into grouped variables selection in a regression, and the adaptive group lasso method is adopted for grouped variable selection. We show theoretically that the new method is able to identify the true number of factors consistently. We also illustrate this model using real and simulated examples. The simulations demonstrate that our approach outperforms some existing methods in terms of the percentage of correct order being detected.

Feature Selection In Segmentation Studies

✦Ewa Nowakowska, GfK Polonia, Smulikowskiego 4, Warszawa, PL-02-785 Poland, *ewa.nowakowska@gfk.com*

Key Words: segmentation, feature selection, clusterability, Fisher discriminant, spectral dimensionality reduction

The process of attribute selection is essential for any segmentation study as it determines the nature of the resulting segments to a great extent. The problem is particularly vital for marketing research where the segmentation solution is frequently the basis for strategic decisions and marketing actions. There exist many different approaches to the topic with the efficiency varied according to the data structure. In this work the main ones are briefly summarized and compared. However, the focus is on a new method that seeks for the set of variables providing the highest distinctness of the clustering structure (clusterability). It is shown, that in the theoretical framework of multivariate Gaussian heteroscedastic mixture model, spectral dimensionality reduction preceded by appropriate data transformation, leads to an easy estimate of the distinctness of the clustering structure. It does not require actual clustering solution which makes it computationally feasible to test and compare clusterability in many different feature subspaces. For illustrative purposes empirical examples are presented and method’s applications for marketing research are also discussed.

Functional Shape Analysis Of Movie Success In Subsequent Sales Channels

✦Yue Tian, University of Maryland, College Park, 4412 Mathematics Building, University of Maryland, College Park, MD 20742-4015 US, *yuetian@math.umd.edu*; Wolfgang Jank, University of Maryland

Key Words: Functional Shape Analysis, interaction term, movies, Word-of-mouth, DVDs

In this study, the goal is to understand (and predict) the shape of the sales curve for entertainment products. By ``shape” we mean whether the curve is trending up over time, or whether it is trending down and the rate at which the shape is changing. Understanding the shape of the sales curve is important for making decisions about production, distribution or marketing. We model the sales curve using functional shape analysis. Functional shape analysis first decomposes the (infinite-dimensional) shape into a set of finite dimensions and subsequently models the individual shape dimensions. In this project, we study the impact of several input shapes on the sales of DVDs over time (i.e. the output shape). The input shapes include sales in prior channels (i.e. at the box office) as well as the evolution of word of mouth about the popularity of a movie. One challenge of our model is the incorporation of an interaction term between the box office shape and the word of mouth shape. We demonstrate how to incorporate interaction terms into functional shape models and compare the resulting model to simpler approaches.

589 Methods for Modelling Weather Data ■●

Section on Statistics and the Environment, Scientific and Public Affairs Advisory Committee

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

A Mixture Probabilistic Model For Extreme Temperatures

✦H. M. Miraz Mahmud, JAHANGIRNAGAR UNIVERSITY, M. S. Student, DEPARTMENT OF STATISTICS, SAVAR, DHAKA, International 1342 BANGLADESH, *miraj_584@yahoo.com*; Mian Arif Shams Adnan, JAHANGIRNAGAR UNIVERSITY; A B M Abdus Sobhan Miah, JAHANGIRNAGAR UNIVERSITY

Key Words: Density Plot, Extreme Value Distribution, Maximum Likelihood Estimation

For various obvious reasons study of extreme temperatures of a particular region is important. In this paper unlike previous studies on extreme temperatures a mixture probabilistic model approach has been applied in case of Jessor Division of Bangladesh. A mixture of extreme value distributions has been found to be appropriate in case of the said region of Bangladesh.

Spatial Nonparametric Approach For Extreme Minimum Temperatures

✦Morgan Lennon, North Carolina State University - Department of Statistics, Raleigh, NC , *morgan_lennon@ncsu.edu*; Montse Fuentes, North Carolina State University; Brian Reich, North Carolina State University

Key Words: spatial statistics, spatial extremes, nonparametric Bayesian, temperature trends

Extreme temperature trends across space and time, especially those of minimum temperature, are not well understood and yet are important for studying climate change. We extend common statistical models for extreme data to the spatial setting using a nonparametric approach by introducing a Dirichelet-type mixture model, with marginals that have generalized extreme value (GEV) distributions. The GEV parameters are allowed to vary spatially and temporally, however this may not explain all spatial correlation. Our proposed methodology is able to capture the unexplained spatial dependence after accounting for the GEV spatial parameters while also allowing for nonstationarity. This modeling approach provides flexibility to characterize complex spatial dependence between the extreme values at the sites. Our approach is computationally efficient, since it avoids matrix inversions used in the common copula frameworks for spatial extremes. The performance of our nonparametric spatial methodology is investigated through a simulation study and applied to minimum temperature data in the Midwest region of the United States.

Statistical Models For Frost Events

✦Reza Hosseini, Simon Fraser University, 8888 University drive, Burnaby, BC V5A 1S6 Canada, *reza1317@gmail.com*; Alireza Hosseini, Yazd University

Key Words: frost, minimum temperature, Agroclimate risk, pistachio

Iran is the largest producer of pistachio in the world (40% of the world's output in 2005). Pistachio trees are very susceptible to spring frost and hence there is a need to assess such risks for agricultural and insurance purposes. This paper develops several statistical models and compare them in terms of frost prediction and model fit. In particular Markov binary frost/no frost models are compared with continuous value models. These models are then used to calculate the probability that a frost happens in a given day or given period.

Long Term Projection On Precipitation Regime Of Thrace Basin Of Turkey Via Nonlinear Time Series Models

✦Sipan Aslan, Middle East Technical University, Department of Statistics, 9006531 Turkey, *sipan@metu.edu.tr*; Ceylan Yozgatligil, Middle East Technical University, Department of Statistics; Inci Batmaz, Middle East Technical University, Department of Statistics

Key Words: Nonlinear Time Series, Precipitation, Forecasting, SETAR, LSTAR, NARX

Thrace basin of Turkey is located between Mediterranean Sea and the Balkan Peninsula, it is surrounded by the Aegean Sea, the Marmara Sea and the Black Sea. Thrace basin known as transition zone of surrounding climate systems for this reason climate of Thrace region and precipitation regime is not consistent. In some years the Black Sea climate or the Mediterranean climate can have an effect on this area. The aim of the study is to clarify unstable precipitation regime and to make a long term projection on it via Nonlinear Autoregressive Models such as Self-Exciting Threshold Auto Regressive (SETAR), Logistic Smooth Transition Autoregressive (LSTAR) and Nonlinear Autoregressive with eXogenous variable (NARX) models. Results of linear parametric time series models such as Seasonal Auto Regressive Integrated Moving Average (SARIMA) and Seasonal Dummy Models are also compared with the nonlinear ones. We used monthly and annual total precipitation observations of land observatories from which are located on Thrace basin of Turkey.

Hidden Markov Models Incorporating Additional Dependence In Regional Rainfall Modelling

✦Nadarajah Iyer Ramesh, University of Greenwich, Department of Mathematical Sciences, Old Royal Naval College, Park Row, Greenwich, London, SE10 9LS United Kingdom, *N.I.Ramesh@gre.ac.uk*

Key Words: Hidden Markov models, Rainfall modelling, Maximum likelihood, Regional precipitation series, Daily rainfall

Hidden Markov models can be modified in several ways to form a rich class of flexible models that are useful in many environmental applications. One of the issues that come up very often when basic hidden Markov models are used to model environmental data is about their ability to accommodate sufficient dependence between observations. We consider a class of hidden Markov models that incorporate additional dependence among observations to model daily rainfall time series. The focus of the study is on models that introduce additional dependence between the state level and the observation level of the process and also on models that incorporate dependence at observation level. Construction of the likelihood function of the models is described along with the usual second order properties of the process. Maximum likelihood method is used to estimate the parameters of the models. Application of the proposed class of models is illustrated in an analysis of regional daily rainfall time series from South East England during 1931 to 2010.

Multisite Precipitation Generator Based On Switching Weather States

✦Oleg Makhnin, New Mexico Institute of Mining and Technology, Math Department, 801 Leroy Pl, Socorro, NM 87801, *olegm@nmt.edu*

Key Words: multisite precipitation generator, multivariate autoregression, Markov switching model, hidden Markov model

In a recent paper, Makhnin and McAlister described a multisite precipitation generator employing multivariate autoregression and truncated and power-transformed (TPT) Normal distribution of the precipitation values. The statistics for the data produced by the generator match the historically observed data fairly well. However, the distribution of dry spells produced by this generator is still not completely satisfying. Here, an extension of the above model is discussed where the current "weather state" is produced according to a Markov switching model, and the multivariate autoregressions depend on the current weather state.

Generalized Linear Mixed Models For Monthly Average Temperatures In A Intertropical Region

✦Mercedes Andrade-Bejarano, Universidad del Valle, Carrera 13 No 100-00, Ciudad Universitaria Melendez, Cali, International 25360 Colomba, *mercedes.andrade@correounivalle.edu.co*; Edinson Zuluaga, Universidad del Valle; John Jairo Millan, Universidad del Valle; Gabriel Conde-Arango, Universidad del Valle

Key Words: Generalised Linear Models, Linear Mixed Models, Isotropic Covariance Function, ENSO phenomenon

Data for this research come from time series of monthly average temperature from 28 sites, collected over the period 1971 to 2002. Due to the geographical location of the study zone (Valle del Cauca, Colombia), monthly average temperature is affected by the altitude and the ENSO phenomenon. Time series for some of the sites show a tendency to increase. Also due to the two dry and wet periods in the study zone, a seasonal behaviour in monthly average temperature is seen. Andrade (2009) modelled monthly average temperature by using linear mixed models, with spatial and temporal correlated errors. However, the errors did not show normal behaviour. The heavy tails of residuals are determined because of the extreme values of temperature in years when the ENSO phenomenon occurs. In this research, we modelled monthly average temperature by using generalized linear mixed models, as an alternative to solve the lack of normality in models fitted by Andrade (2009). We found the Beta distribution as the proper distribution to model the response variable. In the models, years and sites are associated with random effects and altitude, the SOI and the geographical position are fixed effects.

590 Recent Advances in Genome-wide Association Studies

Scientific and Public Affairs Advisory Committee, Section on Statistics in Epidemiology, Biometrics Section, WNAR

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Combining Admixture Mapping And Association Test To Improve The Power Of Genome-Wide Association Studies

✦Jasmin Divers, Department of Biostatistical Sciences, Wake Forest University School of Medicine, WC-2326, Medical Center Blvd., Winston-Salem, NC 27101, *jdivers@wfubmc.edu*; Nicholas M Pajewski, - Department of Biostatistical Sciences, Wake Forest University School of Medicine ; David Redden, University of Alabama at Birmingham; Raymond James Carroll, Texas A&M University; David B. Allison, University of Alabama at Birmingham; Carl D Langefeld, Department of Biostatistical Sciences, Wake Forest University School of Medicine

Key Words: Local admixture, Genome-wide association studies, joint test, measurement error, stistical power, type I error rate

Admixture mapping and genome-wide association studies (GWAS) have been applied separately with various degrees of success in gene mapping efforts. There is evidence that controlling for local ancestry in GWAS may improve the type I error rate in some cases. However, there has been very little emphasis placed on the power of combining admixture mapping with genetic association tests. We have showed in previous work how admixture mapping based on unrelated individuals can be cast as a test of linkage based on identity by descent probabilities. Here we consider various joint tests of linkage and association. We compare their performance under various admixture and population genetic models in an effort to describe cases where such an approach might useful. We also compare the magnitude of the measurement error that occurs in the estimation of the local admixture to the levels observed with the global admixture estimate, and assess their respective effect on these joint tests. We evaluate the effect of the window size selected to obtain the local admixture on the measurement error and the performance of each test.

Regression Based Multi-Marker Tests For Gene-Based Analysis Of Genetic Association

✦Yun Joo Yoo, Seoul National University, , *yyoo@snu.ac.kr*; Shelley B. Bull, Samuel Lunenfeld Research Institute; Lei Sun, University of Toronto, Dalla Lana School of Public Health

Key Words: Global test, Genetic association, Linear combination test, Multiple regression, Linkage disequilbrium, Gene-based test

Single-marker analysis of quantitative traits is predominant in practice, but joint analysis of multiple SNPs can improve power. We propose multi-marker tests called multi-bin linear combination (MLC) tests constructed using parameter estimates from regression of multiple gene-specific tag-SNP markers, with weights determined by the covariance matrix and bins determined by correlation structure of markers. Using power computations and simulations, we compare them with a multi-df joint (Hotelling) test, minimum P tests, tests using principal components (PC), and one-df linear combination tests (LC) under genetic association models specified according to international HapMap genotypes and one or more causal loci for a quantitative phenotype. The LC tests can be more powerful than other tests depending on relationships between causal loci and markers used in the analysis, but suffer loss of power when estimated marker effects have opposite direction, and are sensitive to the marker allele coding scheme. In constrast, MLC tests demonstrated overall stable power gain with improved robustness to the allele coding scheme.

A New Class Of Test Statistics For Detecting Rare Disease Associated Genetic Variants

✦John Ferguson, Department of Epidemiology and Public Health, Yale University, 60 College Street, (Department of Epidemiology and Public Health), New Haven, CT 06520-8034, *john.ferguson@yale.edu*; Hongyu Zhao, Yale University; Judy H Cho, Yale University; Joshua Sampson, NCI/NIH

Key Words: rare variant, association, Score statistic, Data adaptive, GWAS, exon sequencing

For many complex genetic diseases, the common disease associated genetic variation has largely been identified via large scale Genome Wide Association Studies. Unfortunately, rare variants that are disease associated are more difficult to find since the number of individuals carrying the rare allele, which is proportional to the effective sample size in a case control study, will be relatively small when the allele is rare. Recently, many techniques that collectively analyze multiple rare variants within a genomic region have been proposed in an effort to increase statistical power. Here we present a new class of test statistics designed for rare variant association analysis that unifies many of these existing approaches under a common framework. In particular, this class of test statistic encompasses and generalizes a large variety of both uni-variate and multivariate Score tests. Data-adaptive techniques are suggested for choosing a particular member of this class in real data situations. We demonstrate the application of this methodology to find Crohn's disease associated regions using data from a special rare-variant enriched genotyping platform.

A Hidden Markov Model For Haplotype Inference For Present-Absent Genotype Data Using Previously Identified Haplotype And Haplotype Patterns

Jihua Wu, Department of Biostatistics, University of Alabama at Birmingham; Guo-Bo Chen, Section on Statistical Genetics, Dept. of Biostatistics, Univ. of Alabama at Birmingham; Degui Zhi, Department of Biostatistics, University of Alabama at Birmingham; Nianjun Liu, Department of Biostatistics, University of Alabama at Birmingham; ✦Kui Zhang, University of Alabama at Birmingham, RPHB 317H, 1665 University Boulevard, Birmingham, AL 35294-0022, *KZhang@ms.soph.uab.edu*

Key Words: haplotype, hidden markov model, genotype

Killer immunoglobulin-like receptor (KIR) genes vary considerably in their presence or absence on a specific regional haplotype. Because presence or absence of these genes is largely detected using locus-specific genotyping technology, the distinction between homozygosity and hemizygosity is often ambiguous. The performance of methods for haplotype inference for KIR genes may be compromised due to the large portion of ambiguous data. At the same time, many haplotypes or partial haplotype patterns have been previously identified and can be incorporated to facilitate haplotype inference for unphased genotype data. To accommodate the increased ambiguity of present-absent genotyping of KIR genes, we developed a hidden Markov model, which incorporated information about identified haplotypes or partial haplotype patterns and compared several measures on simulated KIR genotype in order to evaluate the reliability of haplotype assignments and the accuracy in estimating haplotype frequency. The simulation study shows that our method outperformed the two existing techniques (HAPLO-IHP and PHASE).

Multi-Layer Clustering And Its Applications: Local Ancestry Inference And Haplotype Based Association Mapping

✦Yongtao Guan, Baylor College of Medicine, 1100 Bates Street Room 2070, Mail Stop 320, Houston, TX 77030, *yongtaog@bcm.edu*

Key Words: local ancestry, admixture mapping, haplotype association, hidden Markov model

We present a multi-layer clustering model that aims to emulate a coalescent process with recombination, where a marginal coalescent tree is approximated by a hierarchy of star trees, and ancestral recombination is approximated by switch of clusters within each hierarchy. In the simplest form of a two layer clustering model, depending on the prior for the cluster-switching frequencies, the bottom layer clustering may represent fine haplotypes and the top layer clustering may represent the structure of such haplotypes, either due to the difference in ancestry or due to subtle difference in haplotype frequencies in case/control status. In the former scenario, one can accurately infer local ancestries for admixed individuals; and in the later scenario, one can detect haplotype based genotype-phenotype association. The model is easy to fit with standard hidden Markov model techniques, can accommodate more than two source populations, and does not rely on training data with known haplotypes.

Multi-Marker Test For Genetic Association

✦LINGZHI LI, UNIVERSITY OF MINNESOTA, 1000 27th Ave SE Apt F, Minneapolis, MN 55414, *lixx1036@umn.edu*; WEIHUA GUAN, UNIVERSITY OF MINNESOTA

Key Words: genetic association, multi-marker test, stepwise selection

We consider the association test between a phenotype and multiple single nucleotide polymorphisms (SNPs) within a gene or region. Due to the low effect sizes of causal variants for many complex diseases, single-locus association tests may not have sufficient power to detect such disease predisposing variants. Here we proposed a new multi-locus approach that sums the genotype scores of a subset of SNPs selected by stepwise selection method, and assesses the significance of summed score through permutation test. The new approach focuses on the SNPs that are more likely to be associated with the disease than other SNPs in the region, and selects the SNPs based on multi-locus regression rather than single-locus tests. We compare our approach with the sum test suggested by Pan et al. [2010], which combines all markers in the region for disease-marker association. The simulation results show that our proposed approach can achieve greater power than the sum test while still maintains correct false-positive rate. We apply our method a multi-center cohort study to identify genes predisposing to different endpoints of kidney transplantation. We hope our method can facilitate the identificati

Power Gains Using Phenotyped But Ungenotyped Relatives In Genetic Association Studies Of Dichotomous Traits

✦Wei Vivian Zhuang, Boston University, , *statinfo88@gmail.com*; Joanne M Murabito, Framingham Heart Study; Kathryn Lunetta, School of Public Health, Boston University

Key Words: Missing Data, Statistical Genetics, Association Studies

In longitudinal studies, some subjects with phenotype data die before providing DNA, but genotyped relatives can be used to impute missing genotypes. We use weighted least square and meta-analysis to explore the power gain for including subjects with missing genotypes in binary-trait-SNP association tests for the case of one SNP and type of relative. The non-centrality-parameter ratio for the association tests including versus excluding ungenotyped subjects depends on the phenotypic correlation and relationship coefficient between the genotyped and ungenotyped relatives, the allele frequency and the variance of the trait given the SNP. Simulations for pedigrees of 2 parents and 2 children using Linear Mixed Effect and Generalized Estimating Equations models yield consistent findings. While the median power ratio increases, the power for a study may decrease when ungenotyped subjects are included. As the proportion of variance explained by the SNP decreases, the probability that the test statistic is decreased increases.

For a dichotomous trait, the inclusion of the ungenotyped but phenotyped subjects improves power in most situations and does not inflate type I error rates.

591 Statistical Methodology in Epidemiologic Reference

Section on Statistics in Epidemiology

Wednesday, August 3, 2:00 p.m.–3:50 p.m.

Applications Of Time Series Methods For Obtaining Complete Mortality Counts

◆Diba Khan, CDC/NCHS, 3311 Toledo Rd, Hyattsville, MD 20176 USA, *ild1@cdc.gov*; Myron J Katzoff, CDC/National Center for Health Sataistics

Key Words: mortality, correlation, forecasting, nonlinear models

The National Center for Health Statistics collects state level mortality data for all causes of deaths through the National Vital Statistics System. The data used in our analysis are weekly mortality counts that cover a span of 8 years for a specific cause of death. Several methods exist in the literature for estimating complete mortality counts. Techniques for modeling and predicting mortality time series are compared when analyzing weekly mortality data. The non-Gaussian structure of the data is taken into account in all analyses. For univariate data, results from the weighted least squares and the quasilikelihood methods are compared. A multivariate approach that exploits the correlation structures for adjoining states is applied to the data for several states within a region. Results are compared when the correlation structures are explicitly modeled in the multivariate approach and when they are excluded from models used in the univariate approach.

Threshold Modeling In The Presence Of Measurement Error: Application To Serum Vitamin D

◆Maria LaVonne Joseph, Iowa State University, 2219 Snedecor Hall, Ames, IA 50011, *josepall@iastate.edu*; Alicia Carriquiry, Iowa State University; Wayne Fuller, Iowa State University; Chris T Sempos, National Institutes of Health

Key Words: measurement error, segmented regression, parathyroid hormone

Vitamin D is essential to ensure bone hardness and strength. Consequently, insufficient levels of vitamin D can lead to undesirable outcomes, such as osteoporosis. Those exhibiting vitamin D deficiency have been characterized by low levels of serum vitamin D (25(OH) D) and elevated levels of serum parathyroid hormone (PTH). A function that has been suggested for the relationship between \textit{usual} PTH and \textit{usual} 25(OH)D, where \textit{usual} represents the long run average of the daily measurements of these quantities, is the segmented regression function. Simulation is used to assess the effects of measurement error on the estimates of the parameters of this function. An estimation procedure based on likelihood methodology is proposed and yields encouraging results in the simulation study.

A Modeling Strategy For Future Survival Days Given A Follow-Up Doubly Censored Sample Of Liver Cancer Patients

◆Vincy Samuel, Florida International University, Department of Epidemiology & Biostatistics, Robert Stempel College of Public Health, Miami, FL 33199 USA, *vsamu002@fiu.edu*; Hafiz M. R. Khan, Florida International University

Key Words: Doubly censored sample, Asymmetrical model, Likelihood function, Highest predictive density interval, Bayesian method, Predictive inference

This talk will focus on a novel application of Bayesian method to estimate future survival days given a set of doubly censored survival days of liver cancer patients. The derivation of predictive densities for future survival days from an asymmetrical model with a Bayesian frame-work will be discussed. A real data example of hepatocellular carcinoma of liver cancer patients who were diagnosed at a local hospital is used to illustrate the predictive results. Furthermore, subgroup levels of the male and female patients' follow-up survival days are considered to obtain the future survival days. The mean, standard deviation, and 95% highest predictive density interval of the future survival days are obtained by making use of the Bayesian method.

Impact Of Early Detection In Prostate Cancer Incidence: A Joint Modeling Approach Of Stage And Age At Diagnosis

◆Chen Hu, University of Michigan, Ann Arbor, Department of Biostatistics, 1415 Washington Heights, Ann Arbor, MI 48109-2029, *chenhu@umich.edu*; Alexander Tsodikov, University of Michigan

Key Words: Prostate cancer incidence, Cancer epidemiology, Screening, Semiparametric, SEER, Disease natural history

Early detection of prostate cancer, such as prostate-specific antigen (PSA) screening, led to a sharp spike in incidence and accompanied by equally pronounced improvement in patient prognoses at diagnosis. Such observation questions the casual connection between early detection and mortality reduction due to the so-called "over-diagnosed" cancers. The impacts of screening on the disease natural history can be justified through statistical models of prostate cancer incidence based on lead time, over-diagnosis and disease progression. In this paper, we propose a semiparametric population model for cancer stage progression and diagnosis to evaluate impacts of screening programs. The disease natural history model is constructed through a series of semiparametric regression with time-dependent covariates such that we can jointly model the correlated response of stage and age at diagnosis as stage-specific cancer incidences. The proposed model provides estimates of lead time, fractions of over-diagnosis and disease progression changes due to early detection. The proposed model is applied on SEER prostate cancer incidence data from 1977 to 2000.

Validating Of Mixture Cure Generalized Gamma Model For Localized Melanoma Prognosis And Outcome

◆Shouluan Ding, Comprehensive Cancer Centor, University of Alabama at Birmingham, VH L108, Box 301, 1670 University Blvd, Birmingham, AL 35294, *sding@uab.edu*; Seng-jaw

Soong, Comprehensive Cancer Centor, University of Alabama at birmingham; Charles M. Balch, Surgery, Oncology and dermatology, Johns Hopkins Medicne

Key Words: Mixture Cure Generalized Gamma Model, Model Validation, Concordance Correlation Coefficient, Melanoma Survival Data

Mixture cure survival model validation is possibly one of the most important and overlooked steps in the complex process of predicting melanoma survival rates. Assessing the predictive power of a given prognostic factor for mixture cure generalized gamma (GG) model with censored data involves multiple challenges. For example, because a model has a statistically significant predictor it doesn't mean that it offers reliable and accurate predictive ability. First, the survival time cannot be adequately predicted due to the censored data. Second, there seems to be no commonly agreed statistical method, especially when a high percentage of data has been censored. Therefore, we propose to borrow the Concordance correlation coefficient (CCC) method to assess the agreement of survival rates at point t between a developed survival model based on a training dataset, and the Kaplan-Meier survival rate based on a testing independent dataset for model validation. The CCC of 2-, 3-, 5- and 10-year for mixture cure GG model were all excellent, at 0.92, 0.93, 0.91 and 0.91, respectively. In each 7 thickness groups, the "cured rate" were estimated at 94%, 86%, 75%, 55%, 49% and 28%, respectively.

Uncertainty And Standard Errors In Agent-Based Disease Projections

✦Georgiy Bobashev, RTI International, 3040 Cornwallis Rd., Durham, NC 27709, *bobashev@rti.org*; R. Joey Morris, RTI International

Key Words: uncertainty, agent-based models, longitudinal, cross-sectional, prediction, behavior

Agent-based Models (ABMs) can be used to quantify future risks by projecting observable behavior into the future. This can be achieved by simulating a hypothetical longitudinal study based on cross-sectional data and estimating quantities on dynamic risks (e.g., relative hazard). Such an approach, however, requires assessment of the variation of the estimates, which would naturally have a higher variance than would be achieved in a real longitudinal study. We present methodology that considers rigorous statistical measurements such as standard errors and uncertainty associated with the fact that the analyzed longitudinal data are a projection of the cross-sectional survey. We illustrate the use of our approach in simulated and real studies.

Mortality And Severe Obesity: Curiously Low Mortality Among The Heaviest Nhanes Iii Participants

✦Scott Keith, Thomas Jefferson University, Philadelphia, PA 19107, *scott.keith@jefferson.edu*; Kevin Fontaine, Johns Hopkins University; David B. Allison, University of Alabama at Birmingham

Key Words: survival analysis, obesity, mortality, BMI, selection bias

The Third National Health and Nutrition Examination Survey (NHANES III; sampled from 1988-1992) has been used to estimate the relationship between measured body mass index (BMI; weight in kilograms/squared height in meters) and mortality in the United States. Although BMI random variables are ratios of continuous quantities, their values are commonly categorized for analysis as underweight (< 18.5), normal weight (18.5-24.9), overweight (25.0-29.9), obese (30.0-34.9), or severely obese (=35.0). While the other BMI categories represent ranges of ~5 BMI units, the severely obese represents 30 units or more. We used various functional bases and BMI categorizations in Cox proportional hazards models, as well as other analytical tools, to examine and contrast mortality rates among the severely obese adults of NHANES III. Preliminary findings suggest that, compared to normal weight, subgroups of the severely obese have surprisingly low mortality rates through 2000 (e.g., HR=1.11 95% CI: 0.69, 1.77 for women with 39=BMI< 45; HR=0.30 95% CI: 0.12, 0.76 for men with BMI>41). This finding could be due to a number of factors and we hypothesize a peculiar form of selection bias.

592 COPSS Awards and Fisher Lecture

ASA, Committee of Presidents of Statistical Societies, International Chinese Statistical Association, SSC

Wednesday, August 3, 4:00 p.m.–5:50 p.m.

Post-Fisherian Experimentation: From Physical To Virtual

✦C. F. Jeff Wu, School of Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, NW, Atlanta, GA 30332-0205, *jeffwu@isye.gatech.edu*

Key Words: computer experiments, experimental design, kriging, factorial experiments, robust parameter design

Experimental design has been a scientific discipline since the founding work of Fisher. During the 80-year history, its development has been largely dominated by work in physical experiments. With advances in high-performance computing and numerical modeling, virtual experiments on a computer have become viable. This talk will highlight some major developments (physical and virtual) in this long period. Fisher's principles (replication, randomization, blocking) will be reviewed, together with principles (effect hierarchy, sparsity, heredity) for factorial experiments. A fresh look at interactions and effect aliasing will be provided, with some surprisingly new insights on an age-old problem. Robust parameter design, another significant development which focuses on variation modeling and reduction, will be mentioned. Turning to computer experiments, the key differences with physical experiments will be highlighted. These include the lack of replication errors which entails new governing principles other than Fisher's and the use of space-filling designs instead of fractional factorials. There are two strategies for modeling and analysis: based on Gaussian processes or on function approximations. These seemingly conflicting approaches can be better linked by bringing a stochastic structure to the numerical errors. Throughout the talk, real experiments/data, ranging from manufacturing to nano technology, will be used for illustration.