



A newsletter published three times a year
by the American Statistical Association - National Council of Teachers of Mathematics
Joint Committee on the Curriculum in Statistics and Probability.

AGAINST ALL ODDS

Three additional programs from the series *Against All Odds: Inside Statistics* have been made available for preview. These programs continue the fine quality of the programs initially reviewed. Their cumulative effect is to make statistics a most appealing field in which to work, and to demonstrate that statistical techniques are the natural methods for exploring important questions.

Program # 11: Causation deals in depth with the relationship of smoking to lung cancer. The inclusion of interviews with key role players in the studies, responses from the tobacco industry, film clips of people smoking (and their resultant coughing), and cigarette advertisements from several time periods all combine to give the program the flavor of a high-quality documentary. The importance of retrospective studies, prospective studies, and laboratory studies, all lending evidence about the causation of lung cancer by smoking, is most impressive. In summary, the arguments for the conclusion of causality were: (1) the consistency of the association in the various studies; (2) the strength of the association; (3) the specificity of the association, including the dose-response effect; (4) the temporal relationship of the associa-

tion (smoking precedes the development of cancer); (5) the coherence of the relationship.

Throughout the program, the viewer is shown clearly that strong association of correlation in itself is not evidence of causation. Examples include the strong relationships between: (1) teacher salaries and amount of alcohol sales; (2) ice cream sales at beaches and drowning at the same beaches (related by a common cause); (3) the decrease in traffic deaths when the national speed limit was reduced (the existence of confounding factors); (4) the growth of population in Louisiana while the land mass is decreasing (coincidence). The term "lurking variables" is used to describe such effects.

Simpson's Paradox is illustrated by a fictitious university with a rigorous business school and a "party-time" law school, each of which admits a higher proportion of female applicants than male applicants. However, when the data from the two schools are pooled, a much smaller percentage of females is accepted than males.

If any single program could convince someone of the value and excitement of statistics, this is the program!

Program # 16: Random Variables uses the most recent Los Angeles earthquake, the failure of the space shuttle Challenger, and the distribution of ice cubes in served glasses of Coca-Cola to convey the notions of independence, the definition of random variable, the distinction between discrete and continuous random variables and their density graphs, and the calculation of the mean and standard deviation of random variables. The O-ring failure on the Challenger flight was shown to be partly the result of lower reliability because of the lack of independence between the functioning of the two O-rings in each field joint. The relatively short and stable interval between earthquakes at the tiny town of Parkfield, California (mean of 22 years, standard deviation of 3 years) is shown to lead to valuable work in earthquake prediction for more populated areas.

Program # 18: Sample Mean and Control Charts, is a natural extension of the ideas developed in program # 16. The analysis of casino gambling (particularly roulette) leads to a discussion of sample means, the central limit theorem, and why the small house advantage pays off through the volume of play. The progressive narrowing of the distribution of the sample means as the number of \$1 bets increases from 50 to 1000 to 100,000 is most dramatic. So is the statement of one casino manager: "The smart bettor would bet everything — all he could afford to lose — on a single play and then walk away from the table, win or lose."

The use of Statistical Process Control charts is illustrated through a visit to a potato chip factory, and the continuing analysis and monitoring of the percent of salt. The discussion includes decision rules which alert the machine operators to make adjustments. Color-coding of the paper on which the control charts are drawn (by one, two or three standard deviations from the process mean) provide visual help to the operators.

The program continues with a brief presentation by Dr. W. Edwards Deming, his ideas, his contributions to Japanese industry, and his consequent discovery and use by American manufacturers.

At the conclusion of the program, the narrator indicates that this concludes the study of probability and sets the stage for statistical inference.

-- Albert P. Shulte

Against All Odds consists of 26 half-hour programs and will be available to PBS stations in September, 1989. Since stations may not schedule this series, a call or letter to your local PBS station may be in order. For details on purchasing the series and off-air taping licenses, write the Annenberg/CPB Project, 2040 Alameda Padre Serra, Box 1922, Santa Barbara, CA 93116-1922 or call 1-800-LEARNER.

-- John Kinney

A DATA ANALYSIS ACTIVITY
Marriages vs. Divorces
in the Northern States
1986

| | State | Marriages | Divorces |
|-----|-------|-----------|----------|
| 1. | ME | 12082 | 5573 |
| 2. | NH | 11708 | 4737 |
| 3. | VT | 5607 | 2377 |
| 4. | MA | 40717 | 19958 |
| 5. | RI | 8044 | 3672 |
| 6. | CT | 26257 | 9152 |
| 7. | NY | 174467 | 59542 |
| 8. | NJ | 60740 | 28397 |
| 9. | PA | 88167 | 40065 |
| 10. | OH | 96994 | 53358 |
| 11. | IN | 50514 | |
| 12. | IL | 95797 | 46511 |
| 13. | MI | 77770 | 37654 |
| 14. | WI | 38386 | 16708 |
| 15. | MN | 34964 | 13893 |
| 16. | IA | 23350 | 10206 |
| 17. | MO | 50273 | 25889 |
| 18. | ND | 5164 | 2264 |
| 19. | SD | 7447 | 2585 |
| 20. | NB | 12027 | 6233 |

The above information, taken from the 1988 World Almanac, gives the number of marriages and divorces for the north and central states in 1986.

1. What do the numbers represent? How do you suppose the data were gathered? What piece of information is missing? Will this affect your analysis? Why do you think it is missing?

2. Make a back to back stem and leaf of the data. What is the usual number of divorces per state? of marriages?

3. Make box plots of the two sets of data. Are there any outliers in either set? Why do you sup-

pose this happened? Which set of data has the most variability? How do the number of marriages compare to the number of divorces?

4. Make a scatterplot of the number of marriages and divorces for each state. Would a $y = x$ line be appropriate for this data? Why or why not?

5. Construct a median fit line if it is appropriate. Write the equation of the line and use it to predict the number of divorces if there were 15,000 marriages in a state. How does New York affect the line?

6. Extension. Are the numbers for the marriages and divorces functions of the population of the state? Construct at least one graphical display that will help answer this question.

-- Gail Burrill, Thong Tong, and Karen Prushka

SATELLITE PILOT PROJECT SET IN KENTUCKY SCHOOLS

On January 17 teacher Tom Graviss met his Probability and Statistics students.

What makes Graviss's students different from other high school math students? They are located at 24 public high schools across the Commonwealth; together, they make up the student body of KET's and Kentucky's first satellite classroom. After more than two years of planning, the KET Star Channels have begun limited operation.

This pilot semester for Graviss and his students, however, reaches beyond Kentucky. Graviss also will teach - via satellite - students at 41 high schools in 16 additional states. This Star Schools Project is sponsored by the Satellite Educational Resources Consortium (SERC), a partnership among 18 state departments of education and educational television networks and stations. KET's production of Probability and Statistics is one of the two courses SERC will offer in its pilot year.

"The KET Star Channels will significantly expand and improve KET's educational services to all Kentuckians," said KET Deputy Executive Director Sandra H. Welch. "We want to make math and science courses, and eventually courses in foreign language, easily accessible to all students in

the Commonwealth. Although it will be three more years before KET's system is totally operational, these first years are very exciting ones for not only the pilot schools but all Kentucky schools.

"It is our goal that the KET Star Channels have a significant equalizing effect on education in Kentucky, making advanced courses taught by outstanding teachers available where it has never been possible before," Welch added.

How will the KET satellite classroom work? According to Welch, KET's distance education, which will link teacher to student through telecommunications technology, is outlined as follows: The daily lesson is taught in the television studio, uplinked from the KET Network Center to the satellite and downlinked to the schools across the state; the students view the program and interact with the teacher using their keypad devices and telephone; the data from the students' keypads and telephone is delivered back to the KET studio via telephone lines and the teacher monitors the students responses (voice and data) and modifies instruction accordingly.

High school students enrolled in the Probability and Statistics course at the pilot schools will receive one-half credit in math. The course will be taught January through May. Teachers in those schools also will participate in interactive training programs produced in Wisconsin and South Carolina.

-- Judy A. Tabor

SOFTWARE FOR USE IN TEACHING STATISTICS

Title: BASIC STATISTICAL PROGRAMS
 Publisher: Frank W. Carlborg
 P.O. Box 1284
 St. Charles, IL 60174
 Hardware: Apple II series of computers with 128K of memory (an exception is noted below), two disk drives and a printer.
 Cost: \$50 (\$55 for overseas shipment)

BASIC STATISTICAL PROGRAMS consists of seven sets of statistical programs (CONCEPTS,

GRAFFIT, ANALYSES, ANOVA, BLOCKS, FACTOR, and RELIABILITY), all of which require 128K of memory with the exception of CONCEPTS which only requires 64K. Each program consists of numerous modules which are menu driven in a user friendly environment. Because the introduction to each module contains a detailed explanation of what is accomplished by the module and what information is required of the user, the programs are self-contained and need no additional documentation. The user generally has the option of printing out this information and other output as needed. The programs most likely to be used in the high school environment are CONCEPTS, GRAFFIT, and ANALYSES and so these are reported on here. The other programs cover more advanced topics in statistics, although some of the modules in ANOVA and BLOCKS might be used in this setting, as well.

CONCEPTS. This program, which is on two disks, develops some of the basic ideas of statistics in an intuitive, interactive manner, making heavy use of graphics and simulation. Some of the topics included are the binomial experiment, p-values, estimation, hypothesis testing, multiple comparisons, regression, and the Central Limit Theorem. Examples from professional football, quality control, and elsewhere are integrated throughout. Although this reviewer likes the aim and scope of this program, he finds it unacceptable in its current form for high school instruction for the following reasons.

1. The author defines terms in a nonstandard way. For example, if an experiment is to flip a coin 16 times, then he refers to $(0, 1, \dots, 16)$ as the "space" of the experiment and (4) is the event that 4 heads are observed. Of course, the sample space of the experiment is a set of sequences of heads and tails (there are 2^{16} of them) and if x , say, is the number of heads in a particular sequence, then the range of (the random variable) x is his "space." Also, the definition of p-value is unusual, leading to incorrect p-values in many cases.

2. The author refers to numbers "between A and B," but means "from A to B."

3. The method for deriving the p-value for a set of multiple comparisons is arbitrary and nonstandard.

4. "Best estimates" are given in the regression problem with no explanation of what they are or

why they are "best."

5. Confidence interval estimation is treated as though it is a separate topic from hypothesis testing and as if no hypothesis can be tested, no decision can be made, and no calculation of a p-value is possible from a confidence interval.

6. The graphics for the Central Limit Theorem are great, but the program is too slow.

7. A more careful organization of modules and topics within modules would be appropriate. Right now, if a user proceeds sequentially through the modules, p-values will be encountered, for example, before they are explained or defined.

8. It is possible to be dumped from the program by typing a response other than one of those expected by the computer.

9. The time spent moving from module to module and within a module is exasperatingly slow. One of the available speeded-up versions of DOS would be a useful addition to this program as well as the others.

GRAFFIT. This program, which is on two disks, permits the user to generate high resolution graphs, including up to 6 side-by-side box and whisker graphs, relative frequency histograms, probability plots, 2 and 3 dimensional scatter diagrams, plots of functions of the form $f(x) = a + bx + cx^2 + dx^3$ or $f(x, y) = a + bx + cy + dxy + ex^2 + fy^2$, where the user specifies the coefficients and the range of x and y values, time series plots, autocorrelation and cross-correlation graphs, and appropriate graphs for regression analysis. Included is a data management system which enables the user to enter, transform, and store data for use in this and the other programs. The graphs obtained by this program are excellent with the only deficiency being that the axes, which are calibrated according to user specification, are not labeled. It would be nice to have line plots and stem-and-leaf plots available, also.

ANALYSES. This program, which is on a single disk, permits the user to construct confidence interval estimates and conduct tests of hypotheses in one-sample, two-sample, one-way ANOVA, two-way ANOVA, randomized block designs, and about one or more parameters in a linear regression model. The same data management system that is included with GRAFFIT is on this disk, en-

abling the user to enter, transform, and store data for this program and other programs. Although the output from the various modules is easy to read and understand, there are enough deficiencies to make this program unacceptable in its present form for use in a high school setting. In many modules, output is sent directly to a printer and never appears on the screen. The user should have the option of a screen display as well as a printout. In the randomized blocks module and elsewhere, "treatments" are confusingly called "samples." Other, similarly abusive use of terms occurs in other modules and frequently enough to possibly be confusing to beginning statistics students who are familiar with standard terminology that has been carefully defined.

SUMMARY. The idea behind this set of programs is excellent — to provide high school students with an inexpensive, self-contained, quality statistics package for the Apple II series of computers which provides excellent graphics and makes number crunching easy. In many respects, this package comes close to achieving this goal, particularly, the programs in GRAFFIT. Unfortunately, the other two programs considered in this review were sufficiently marred by imprecise and vague terminology and a slow operating system to make them generally unacceptable for use in the high school setting. An upgraded version of this software might be much more highly rated.

— James L. Kepner

QUANTITATIVE LITERACY: THE SERIES AND THE SOFTWARE

The four curriculum units which have been written under the auspices of the ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability and the Quantitative Literacy (QL) Project are now available in both student and teacher editions from:

Dale Seymour Publications
P.O. Box 10888
Palo Alto, CA 94303

If you would like more detailed information about the units, there is an excellent *Quantitative Literacy Series Program Sampler* which also may

be ordered at no cost from Dale Seymour at the address above.

QL Software

The diskette designed to accompany the QL Series is also complete, thanks to the efforts of many people around the country. It is available in both Apple II series and IBM versions. The programs are partitioned into four modules with each module designed to implement one of the units in the series. The programs are menu driven, user friendly, and designed to be used in conjunction with the units with minimal effort.

The first module contains programs which generate box plots, stem-and-leaf plots, scatter plots, and fit a line to bivariate data on scatter plots.

Except for the stem-and-leaf program, which requires the user to supply his own data, all of the programs allow the user the option of using the pre-programmed data from the applications in *Exploring Data* or entering his own data.

The second module of programs was designed to supplement the unit *Exploring Probability*. Included are numerous programs which generate outcomes from dice rolling, coin flipping, and random digit selecting experiments.

The programs in the third module were designed to accompany *The Art and Techniques of Simulation*. The first program in this section leads a student through the 8 steps involved in setting up a simulation, and each of the remaining programs demonstrates a particular simulation application. Extensive graphics are available at the option of the user. The cost of the diskette is \$2 (with a disk) or \$7 (if a disk is not sent) for either the IBM or Apple version. Write Gail Burrill, Whitnall High School, 500 S. 116th St., Greenfield, WI 53228.

BOOKLETS ON STATISTICS

Careers in Statistics

This 24-page booklet introduces the mathematically inclined student to the world of statistics. Some of the topics covered are What Do Statisticians Do?, Educational Training and Where to Begin Your Statistical Career. Also included is a list of the U.S. and Canadian schools that offer degrees in statistics.

There is no charge for orders of 25 or less; additional copies are 25¢ each and must be prepaid.

Minorities!...Looking for a Challenging Career?

Blacks, Hispanics, American Indians and Asian Americans are insufficiently represented in the field of statistics. Special efforts are being made to inform minorities of advantages and opportunities for careers in statistics. The brochure discusses the field of statistics, job opportunities, program of study and salaries.

There is no charge for orders of 25 or less; additional copies are 10¢ each and must be prepaid.

What Is A Survey?

What Is A Survey? is interesting reading for the teacher or the secondary student. This 32-page booklet is written for the non-statistician. It contains information on types and characteristics of surveys, designing and conducting a survey, sources of errors, and using the results of the survey.

Single copies are free; additional copies are 50¢ each and must be prepaid.

Statistics as a Career: Women at Work

This brochure is designed to encourage women to enter the field of statistics. It includes a job description and picture of six women statisticians in government, research, teaching, and industry. It also addresses "The Demand for Statisticians," "How Much Do Statisticians Earn?," and "How to Become a Statistician."

There is no charge for orders of 25 or less; additional copies are 10¢ each and must be prepaid.

Statistical Science: The Profession

Introduces senior high school students to statistics as a science and a profession by using a slide presentation or VHS tape. The presentation introduces the field of statistics and shows the diversity of applications of statistics to problems in our technological society. The slide package includes 80 professionally developed slides, an accompanying script, and instructions.

There are six modules in the package: an introduction; examples; educational requirements and job opportunities, applications, a summary, and a finale. The presentation is approximately 25 minutes.

The cost of the slide presentation is \$65; the cost of the VHS tape is \$50. Each may be rented at a rate of \$10 per week which can be applied toward a purchase. All orders must be prepaid. Order from the American Statistical Association, 1429 Duke St., Alexandria, VA 22314-3402.

NEW PUBLICATIONS AND PRODUCTS

Proceedings of the Second Conference on the Teaching of Statistics

Nouri, Esmat (Editor)

Department of Mathematical Sciences
State University of New York
Oneonta, NY 13820-1368
1988, Paperbound, \$20

This volume consists of a compilation of the papers and discussions presented at the Second Conference on the Teaching of Statistics held at the State University of New York College at Oneonta on April 24-25, 1987. The objective of the conference was to promote statistical education. There were two key goals: to promote the interchange of ideas about teaching material and to foster cooperation between educational institutions and the government agencies and industries that use statistics.

A wide range of issues was addressed by approximately 100 participants from fifteen states representing high schools, colleges and universities, government agencies, and private industries. Topics discussed included content and method for an introductory statistics course, recommendations for statistics curriculum in high schools, career opportunities and statistical training for government and industry, and microcomputer statistics packages. Two papers are briefly highlighted here to illustrate the concerns of participants and to provide a flavor of the proceedings.

One interesting and informative paper which is presented in the text is a panel discussion on the statistics curriculum in high schools. The discussion centers around the Quantitative Literacy Project and the curriculum materials developed by

members of the ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability. The paper provides an excellent overview of the QLP and the QL Series for those who are unfamiliar with the recent efforts of the Joint Committee.

Another paper worthy of attention is the presentation from the panel discussion on career opportunities and statistical training for government and industry. Panelists included representatives from the U.S. Department of Agriculture, IIT Research Institute, AT&T Bell Laboratories, Abbott Laboratories, The Bureau of the Census, and Corning Glass Works. The importance of statistics in quality and productivity in industry was one major focus of the discussion which included specific recommendations from panelists concerning qualifications for statisticians seeking positions in industry and government.

Contact the editor, Esmat Nouri, to order the book.

-- Beth Bryan

BOOK REVIEW

Probability and Statistics. Dale M. Johnson. 1989. ix + 453 pp., \$14.25. ISBN 0-538-60015-2. South-Western Publishing Company, 5101 Madison Road, Cincinnati, OH 45227.

This readable text is designed for a semester course in probability and statistics. Separated into four major divisions - Descriptive Statistics, Probability, Inferential Statistics Concepts, and Inferential Statistical Methodology - the book treats topics found in traditional statistics courses.

Each chapter ends with a Computer Focus section whose purpose is to teach students to write BASIC programs to solve statistics problems. A computer disk, ACCUSTAT, is available to provide answers to selected chapter applications exercises identified in the text by a computer icon. A site license for ACCUTSTAT is available.

Applications and examples from the field of education are used liberally throughout the text. According to the author, this practice enables students to concentrate on the essentials of probability and statistics rather than trying to understand an unfamiliar field before attacking the statistical issues involved. Because of this approach, students

see few applications of statistics to areas other than education.

The chapters are relatively short with problem sets appearing at the end, rather than interspersed throughout the chapters. Each chapter contains an average of fourteen problems. Of these, an average of seven are marked for computer solution.

Minor attention is given to the core concepts of the ASA/NCTM Quantitative Literacy Project. The short sections on stem-and-leaf plots, box-and-whisker plots, and median fit lines are labelled as optional. Simulation of probabilities is not discussed.

Supplementary materials include a Study Guide containing practice problems and a Teacher's Resource Guide containing teaching suggestions, problem solving activities, transparency masters, a copy of the Study Guide with answers, and chapter tests and answers.

-- Pamela Coffield

TEACHING STATISTICS

The journal *Teaching Statistics* is now ten years old and has recently undergone a change of format. Its purpose remains: to aid those who teach statistics as a subject in its own right or as part of a mathematics course; to help teachers of geography, history, economics, social science and science by showing how statistical ideas can illuminate their work; and to encourage the proper uses of statistics in teaching.

Regular features of the journal include Project Parade, Practical Activities, Historical Perspective, Computing Corner, Curriculum Matters, Data Bank, and Problem Page. A typical issue also includes about six individual articles of interest to teachers of statistics.

Dr. David Green of the University of Technology, Loughborough, U.K., is the new editor.

A one-year subscription (three issues) is \$20/£12.00 (overseas) or £9.50 (U.K.). Subscription inquiries should be addressed to: Longman Group U.K. Ltd., Fourth Avenue, Harlow, Essex, CM195AA U.K.

-- John Kinney