JUNE 1990

WASHINGTON STATISTICAL SOCIETY

NEWSLETTER

June 7	Thursday	Visualization of Hyperdimensional Data Structures
June 14	Thursday	Effects of Procedural Differences in the Nationwide Food Consumption Survey
June 19	Tuesday	Interim Analyses: The Repeated Confidence Interval Approach
June 20	Wednesday	Achieving Comparable State Fiscal Data in Education: The NCES Crosswalk Project

ANNOUNCEMENTS

1991 Annual Research Conference - Call for Papers

The Bureau of the Census is planning its 1991 Annual Research Conference, to be held in March of 1991 in the Washington, DC area. The conference will consist primarily of contributed papers, most of which receive formal discussion at the conference. The conference will feature papers on topics related to a broad range of Census Bureau research interests. Papers may address methodology, empirical studies, or relevant issues. A conference proceedings volume containing all papers and discussions will be published. Papers must be original and not previously published or disseminated. Presenters will be reimbursed for transportation and per diem expenses within the USA and will receive a fee for manuscript preparation (expected range: $250--$450).

Topic areas include:

- Data collection technologies
- Developments in survey processing (databases, editing, etc.)
- Impact of automation on survey organizations
- Coverage improvement in economic censuses and surveys
- Economic measurement methods and issues (housing affordability, well-being, poverty, etc.)
- Labor force concepts and measurement
- International comparisons of survey estimates
- Longitudinal methodology for survey data
- Cross-sectional, retrospective and longitudinal measurement of change
- New statistical uses of administrative data
- Post-censal estimates of small area populations
- Model-based approaches to sampling theory
- Modeling and measuring nonsampling errors
- Collecting sensitive behavioral data
- Dependent vs. independent interviewing
- Effects of interviewers and respondents on data quality
- Cognitive and noncognitive aspects of questionnaire design
- Geographic/cartographic methods for censuses and surveys *(continued on page 4)*

WASHINGTON STATISTICAL SOCIETY PROGRAM CHAIRS

Agriculture & Natural Resources
Cynthia Clark 763-8558
Paul Gargiullo 586-1137

Social & Demographic Statistics
Harvey Schwartz 443-6990
John Czajka 484-9220

Short Courses
Glenn White 763-7524
Donald Gantz 764-6565
Brad Pafford 447-3623
Sid Schwartz 268-3490

Economics
John Ruser 523-1347
Neil Ericsson 452-3709

Methodology
Myron Katzoff 436-7047
Andrew White 436-7904

Public Health & Biostatistics
Ed Lakatos 496-5905
Jai Choi 436-7047

Physical Sciences & Engineering
Patricia Abel 883-6490
Refik Soyer 994-6794

Statistical Computing
Khalid Aboura 994-6794
Sylvia Leaver 272-2350

Quality Assurance
Stanley R. Freedman 586-2038
John Galvin 272-5066

Newsletter Editor
Michael Cohen 454-6193

Employment
Bill Arends 447-6812

PROGRAM ABSTRACTS

TOPIC: **VISUALIZATION OF HYPERDIMENSIONAL DATA STRUCTURES**

SPEAKER: Edward J. Wegman, Center for Computational Statistics, George Mason University

CHAIR: Sylvia Leaver, Bureau of Labor Statistics

DATE & TIME: Thursday, June 7, 1990; 12:30 to 2:00 p.m.

LOCATION: Room 2437, GAO Building, 441 G Street, N.W., Washington, D.C.
(Sign in at guard desk and state purpose and room number of visit.)

SPONSOR: Statistical Computing Section

ABSTRACT: Statistical inference focuses on the development of methodologies for characterizing the probabilistic structure of one or more random variables. This may be done in the context of parametric or nonparametric estimation or hypothesis testing. In contrast, the problem of structural inference focuses on inference about the relationship between (among) two or more random variables. Linear regression and time domain time series modeling are the two classical examples of structural inference. In structural inference problems, graphical displays often play a critical role in determining the model to be posed.

This lecture will focus on two aspects of the problem. We will illustrate some approaches to visualizing higher dimensional data structures. These will principally include parallel coordinate scatter plots and parallel coordinate density plots, but will also compare these to scatter plot matrices, rotating scatter plots, star diagrams and several other visualization tools. In addition, we consider the grand tour introduced by Asimov (1985), which is based on the idea that one method of searching for structure in d-dimensional data is to "look at it from all possible angles," more mathematically, to project the data sequentially into all possible 2-planes. The collection of 2-planes in a d-dimensional space is called a Grassmannian manifold. A key feature of the grand tour is that the projection planes are chosen according to a dense, continuous path through the Grassmannian manifold which yields the visual impression of points moving continuously.

Of course, while the grand tour just described will reveal non-random 2-dimensional structure, it may not be particularly helpful in isolating higher dimensional structure. We propose the k-dimensional grand tour in d dimensions, where k = d. We give basic algorithms for computing a continuous sequence through the Grassmannian manifold of k-flats. We use the k-dimensional parallel coordinate display to represent visually the projections of the data into k-flats.

The use of all these techniques will be illustrated with real multivariate data sets.

TOPIC: **EFFECTS OF PROCEDURAL DIFFERENCES IN THE NATIONWIDE FOOD CONSUMPTION SURVEY**

SPEAKER: Patricia Guenther, Human Nutrition Information Service

CHAIR: Cynthia Clark, Census Bureau

DISCUSSANT: Charles Cowan, Opinion Research

DATE & TIME: Thursday, June 14, 1990; 12:30 to 1:30 p.m.

LOCATION: Room 2437, GAO Building, 441 G Street, N.W., Washington, D.C.
(Sign in at guard desk and state purpose and room number of visit.)

SPONSORS: Agriculture and Natural Resources Section

PROGRAM ABSTRACTS (continued)

ABSTRACT: The 1988 Bridging Study was conducted to facilitate comparison of results from the 1987-88 Nationwide Food Consumption Survey (NFCS) with results from the 1977-78 NFCS. A field experiment was designed using a split-sample approach to test the effects of changes in interview, food coding, and weight conversion procedures and nutrient data bases. Group A was interviewed using 1987 procedures, and their nutrient intakes were calculated using 1987 food codes, weight conversions, and nutrient data base. Group B was interviewed using 1977 procedures, and their nutrient intakes were calculated four ways, using various combinations of 1987 and 1977 food codes, weights, and nutrient data bases. For the most part, effects of the various changes in survey procedures were slight and tended to offset each other.

TOPIC: **INTERIM ANALYSES: THE REPEATED CONFIDENCE INTERVAL APPROACH**

SPEAKER: Bruce W. Turnbull, Department of Operations Research and Industrial Engineering, Cornell University

CHAIR: James Dambrosia, Biometry and Field Studies Branch, NINCDS, National Institutes of Health

DISCUSSANTS: Gordon Lan, Biostatistics Center and Dept. of Statistics, George Washington University
Laurence Freedman, Biometry Branch, NCI, National Institutes of Health

DATE & TIME: Tuesday, June 19, 1990; 1:00 p.m. (Note special time.)

LOCATION: Conference Room 7, Building 31, C Wing, National Institutes of Health
(from Medical Center Metro - take shuttle (every 15 min. on the hour) or left on South Dr., right on Center Dr., - when Center Dr. turns left, Building 31 will be on your right.)

SPONSORS: Biometry and Field Studies Branch, NINCDS, NIH
and Biostatistics and Public Health, WSS

ABSTRACT: Most medical trials are monitored for early evidence of treatment differences or harmful side effects, and many sequential methods have been proposed for this. Similarly, data collected in a prospective epidemiological study are likely to be reviewed periodically in the course of the study. The repeated confidence interval approach, which combines aspects of sequential estimation and testing, allows a full exploration of the data at each interim analysis and does not depend on a rigidly enforced statistical stopping rule. In this paper we present the general principles underlying the construction of repeated confidence intervals and describe how they can be used in reaching a decision to terminate a study early. We discuss design considerations, which depend on the form of early stopping anticipated, and explain how the basic method can be adapted to cope with the problems of unpredictable group sizes or, more generally, unequal increments in information between analyses. Extensions of the method to handle survival data, categorical data, normal responses with unknown variance and multivariate normal observations are also presented.

TOPIC: **ACHIEVING COMPARABLE STATE FISCAL DATA IN EDUCATION: THE NCES CROSSWALK PROJECT**

SPEAKERS: Martin Orland, National Center for Education Statistics
Joel Sherman, Pelavin Associates Inc.

CHAIR: Stan Freedman, Energy Information Administration

DISCUSSANT: Mary Batcher, Internal Revenue Service

DATE & TIME: Wednesday, June 20, 1990; 12:30 to 2:00 p.m.

LOCATION: Room 2736, GAO Building, 441 G Street, N.W., Washington, D.C.
(Sign in at guard desk and state purpose and room number of visit.)

SPONSOR: Quality Assurance Section

PROGRAM ABSTRACTS (continued)

ABSTRACT: For the past three years the National Center for Education Statistics has been developing State fiscal crosswalks to translate disparate State administrative financial records into a common national reporting format. This presentation will describe the nature and status of this effort including the need for the activity, basic project dimensions, technical and substantive barriers encountered in implementing this initiative, and results to date.

ANNOUNCEMENTS (continued)

1991 Annual Research Conference - Call for Papers (continued)

To have a paper considered for presentation, send a 500-word abstract by August 13, 1990 to: **Lynn Weldman, Conference Chair, Statistical Research Division, Bureau of the Census, Washington, D.C. 20233.**

To obtain registration information or to be included on the mailing list, contact: **Maxine Anderson-Brown, Conference Coordinator, Office of the Director, Bureau of the Census, Washington, D.C. 20233.**

Please note, plans for ARC 1991 are dependent upon the final approval and funding which are still pending.

43rd Annual Summer Institute in Survey Research Techniques at ISR

The Survey Research Center of the Institute for Social Research, The University of Michigan, will hold its 43rd annual Summer Institute in Survey Research Techniques at ISR during the summer of 1990. Among the courses being offered, the Summer Institute will sponsor a one-week short course focusing on the Panel Study of Income Dynamics (PSID). Concurrently, there will be a short course on event history methods that can be used to analyze data from the PSID and other longitudinal studies. These courses will each be offered during the week of July 23-27 at the Institute for Social Research in Ann Arbor, Michigan. The time schedule of these workshops will be coordinated so that participants can attend both courses. Of course, participation in either course is not conditional upon taking the other. The fees for the courses are $600 ($400 for members of institutions in the Inter-University Consortium for Political and Social Research). Participation in both workshops is $800 ($600 for ICPSR members).

Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID) is a long-term, large-scale panel study of families and households, with annual data on an initial sample of 5,000 panel households obtained since 1968 and extending to the present. Data now exist on more than 7,000 households and 35,000 individuals. The PSID now has a diverse set of event-history data, some of which extend over the entire twenty-year PSID interviewing period. The course is offered in conjunction with the short course on event-history analysis; however, application of PSID data in analyses not involving event history methods will also be covered in this short course. This course will be taught by Greg Duncan, Martha Hill, and members of the PSID project staff.

Event History Analysis

This short course, taught by Jay Teachman (University of Maryland), focuses on the application of hazard-rate models of analysis to event-history data. The use of event history data, recording the timing and duration of life experiences, is becoming increasingly common in the social sciences. The application of conventional analysis tools to event histories can produce severe bias or loss of information. This course shows how these difficulties can be overcome, enabling the analyst to realize the full benefit of event history data. The emphasis is on "regression-like" models, *(continued on page 5)*

ANNOUNCEMENTS (continued)

43rd Annual Summer Institute in Survey Research Techniques at ISR (continued)

commonly known as proportional hazards models or hazard-rate models, in which the likelihood of an event occurring varies according to a set of explanatory variables. Topics will include: data structure, life tables, censoring, discrete-time methods, parametric models, nonparametric models, time-varying explanatory variables, unobserved heterogeneity, competing risks, and repeatable events. Computer programs useful for the analysis of event history data will be discussed and illustrated using several survey databases, including the PSID.

THE SUMMER INSTITUTE IN SURVEY RESEARCH TECHNIQUES

In addition to the short courses described above, the annual Summer Institute will offer 18 regular 4-week graduate-level courses in various aspects of survey research design, data collection and analysis. The emphasis in the program is on the sample survey as a basic measuring instrument for the social sciences. Teaching faculty in the Summer Institute are drawn primarily from the Survey Research Center and well-known experts in the field of survey research. The courses are offered through the graduate programs of Departments of Sociology and Psychology of the University of Michigan. Participants in the program gain familiarity with the application of survey research methods, including research design, sampling, measurement, questionnaire design, field methods, data management, and the statistical analysis of data. The following is a list of these courses and the instructors. The 1990 Summer Institute will be held in Ann Arbor from July 2 to August 24.

First Session — July 2 through July 27

- Introduction to Survey Research, Jason Lee (Northern Illinois University)
- Introduction to Statistical Research Design, Karl Landis (University of Michigan)
- Analysis of Survey Data, Willard Rodgers (University of Michigan)
- Questionnaire Design, Nora Cate Schaeffer (University of Wisconsin)
- Design of Evaluation Research, Bill Yeaton and Paul Wortman (University of Michigan)
- Mail and Telephone Survey Methods, Don Dillman (Washington State University) and Paul Biemer (New Mexico State University)
- Methods of Survey Sampling, Graham Kalton (University of Michigan)
- Improving the Quality of Survey Data, Lars Lyberg (Statistics Sweden)

Second Session — July 30 through August 24

- Introduction to Survey Research, Jason Lee (Northern Illinois University)
- Advanced Methods of Survey Sampling, Jim Lepkowski (University of Michigan)
- Computer Analysis of Survey Data, Laura Klem (University of Michigan)
- Regression Analysis of Survey Data, Willard Rodgers (University of Michigan)
- Longitudinal Analysis of Survey Data, Duane Alwin (University of Michigan)
- Event History Analysis, Jay Teachman (University of Maryland)
- Cognitive Psychology and Survey Methods, McKee McClendon (University of Akron)
- Research on Survey Quality, Duane Alwin (University of Michigan)

For a detailed brochure for the Summer Institute, contact Dr. Duane F. Alwin, Director of the Summer Institute, Survey Research Center, The Institute for Social Research, The University of Michigan, P.O. Box 1248, Ann Arbor, MI 48106-1248. Telephone (313) 764-6595.

EMPLOYMENT COLUMN

The Washington Statistical Society Newsletter provides a service of notification of employment opportunities and descriptions of those seeking employment here in Washington. Readers are encouraged to take advantage of this feature of the newsletter. Deadline for inserting notices is 5 (five) weeks before the publication date. Those interested should write to: Bill Arends, USDA-NASS, Room 4133 South Building, Washington, D.C. 20250-2000, Phone 447-6812.

JOB OPENINGS

Biostatistician/Epidemiologist

Children's Hospital National Medical Center announces an opening for a Biostatistician/Epidemiologist in the Office of the Research Director. This person will join a team who provide statistical consultation on a variety of research projects for clinical faculty and staff. Responsibilities include:

- Review of protocols
- Design of clinical trials and laboratory experiments
- Analysis of data
- Preparation of reports
- Conduct of seminars

The position requires a Ph.D. in biostatistics, statistics, or epidemiology, with at least 2 years of consulting experience in an academic or medical environment. Experience using statistical software (e.g., BMDP, SAS, SPSS) is essential. Knowledge of a scientific program language (e.g., Fortran) is preferred.

The position offers opportunities for an academic appointment to the George Washington University School of Medicine and for pursuing an individual course of research. For more information contact: **Dr. Pamela Getson at 745-3323** or send application materials to **Nick Piazza, Human Resource Department, 111 Michigan Avenue, N.W., Washington, D.C. 20010.**

GH-13 Mathematical Statistician
Program Research and Development Branch
Agriculture Division
U.S. Census Bureau

The Agriculture Division of the Census Bureau is currently recruiting for a new mathematical statistician position in the Program Research and Development Branch, GH13/1529 ($42,601-$55,381). The incumbent will supervise a staff of three to five mathematical statisticians and work with a group of mathematical and survey statisticians. The branch is responsible for developing a list frame of all agricultural operations, evaluating census coverage, and designing and conducting evaluation surveys, test censuses, and survey processing research for the programs of the division. Specific project responsibilities will include design of the nonresponse survey sample, design of coverage evaluation area and list samples, coverage estimation, census and questionnaire test design and analysis, and agricultural statistical data analysis. Interested applicants should contact **Cynthia Clark, Assistant Division Chief for Research and Methodology, immediately, at (301) 763-8558.**

JOB OPENINGS (continued)

Part-Time Position to Become Available Soon

Wanted: Statistician to serve as Director of Information Services

at the
American Association of Colleges of Osteopathic Medicine
Rockville, Maryland

- Permanent position, approximately 20 to 25 hours per week
- Requires statistical training and experience, computer literacy, and ability to prepare analytic reports with photos and tables.
- Experience with Public Health Service is desired.

Contact: **Don Zobell, 301-468-0990**

STATISTICIAN (Service Fellow-equivalent to GS 11/12/13)

The Agency for Health Care Policy and Research (AHCPR) is seeking a statistician to join the statistical staff of the National Medical Expenditure Study. The position is available for a Ph.D. in statistics/biostatistics or a M.S. with experience in sample design, survey research, sampling, weights development, data analysis for complex surveys, imputation procedures and matching techniques. Familiarity with statistical software packages (SAS, SPSS) is required. Send SF-171 application forms and resume to: **Dr. Steven B. Cohen, Senior Research Manager, Agency for Health Care Policy and Research, Room 18A-55, 5600 Fishers Lane, Rockville, MD 20857,** or call him at (301) 443-4836 for additional information.

JOB APPLICANTS

CODE #90-04

Objective:	Position as Applied Mathematician, specializing in Statistical Analysis Preference: Private Industry in Northern Virginia or D.C. Metropolitan area
Professional Experience:	Seasonal Adjustment of Time Series data (2 years) Econometric Forecasting (2 years) Mathematical Optimization Techniques (3 years) Reliability of Computer Systems (1 year)
Computer Skills:	TSO, JCL, and WYLBUR on IBM mainframe Micro computer communication packages Crosstalk and PcLink SAS (5 years) on mainframe and micro computer
Education:	B.A. in Mathematics With Specialization In Statistics Connecticut State University

JOB APPLICANTS (continued)

CODE #90-05

Objective: Management Engineering. Consulting in Adaptive Process Control and Industrial Statistics. Design and Development of Industrial Control and Information Systems.

Experience: Management Engineering. Three years of experience in management of innovation of control quality systems for metallurgy.

Consulting. Over ten years experience consulting and lecturing for industrial companies and universities.

Design and Development. Fifteen years of experience in designing and developing control systems and new technology.

Publications: Over ninety publications on theoretical and applied statistics and on industrial applications of adaptive control.

Education: Ph.D. in Control Theory and Industrial Statistics from Moscow Steel and Alloy Institute, Moscow, USSR (The Doctor of Science Degree).

Ph.D. in Statistics and Information Theory (The Candidate of Science Degree), and M.S. in Statistics and Information Theory from the Moscow Institute of Physics and Technology, USSR.

Personal: US permanent resident. Full authorization to work in the US.

WASHINGTON STATISTICAL SOCIETY

P.O. Box 70843
Washington, D.C. 20024-0843

NON-PROFIT ORG.
U. S. POSTAGE
PAID
FALLS CHURCH, VA.
PERMIT No. 186