September 1987

WASHINGTON STATISTICAL SOCIETY

# NEWSLETTER

| | | |
|---|---|---|
| September 17 | Thursday | Laspeyres Price Index Estimation Under an Autoregressive Model |
| September 29 | Tuesday | Two-Sample Linear Rank Tests for Censored Data |
| October 7 | Wednesday | Generalized Variance Functions for a Complex Sample Survey |

## ANNOUNCEMENTS

### Classification and Regression Trees

VIDEO SHORT COURSE

A powerful nonlinear alternative to the classical linear procedures for multivariate data analysis is the construction of classification and regression rules in the form of binary decision trees. This method has been used successfully in medical, chemical, financial, and meteorological applications, as well as in other fields.

On Thursday, November 12, from 9 a.m. to 5 p.m. the Washington Statistical Society and George Mason University's Center for Computational Statistics and Probability will co-sponsor a one-day video short course on ***Classification and Regression Trees***, located at George Mason University, Student Union Building #2, Rooms 5-7, Fairfax, Virginia. This is a videotape of the one-day short course presented at the 1985 annual meeting of the American Statistical Association by Drs. Jerome Friedman, Richard Olshen, and Charles Stone. Jerome Friedman, Professor of Statistics at Stanford University and Director of the Computational Research Group at the Stanford Linear Accelerator Center, will be present to introduce each of the course modules, to discuss recent developments and specific issues in greater detail, and to entertain questions from the audience.

Tree-structured rules are easier to interpret than the classical ones, especially in the context of multiple classification. The methodology is flexible, allowing for categorical and continuous predictor variables and for interaction effects of unspecified form. It has the natural ability to select informative variables and disregard non-informative ones automatically. The use of cross-validation leads to the construction of useful trees and yields an honest evaluation of their accuracy where there are only 200 cases in the training sample; but the information in large data sets having tens of thousands of cases can also be used to advantage. This computer-intensive methodology provides an effective tool for gaining insight rapidly into important structural aspects of the data being investigated.

The course will cover this methodology, following the book ***Classification and Regression Trees***, by Breiman, Friedman, Olshen, and Stone, and using CART™ software, developed by the same authors, for illustrative purposes. The course is recommended for statisticians, engineers, scientists, and others whose work involves classification or regression and for teachers and students of statistics. The audience is assumed to have had a course in probability and statistics and some familiarity with classification or regression.

**OUTLINE**

- Introduction to tree-structured classification
- Right-sized trees and their evaluation
- Splitting rules
- Strengthening and interpreting
- Real world applications
- Tree-structured regression

All registration materials, including purchase order, must be received by October 30, 1987. To receive confirmation of your registration, please enclose a self-addressed, stamped envelope with your registration material. Please contact Donald Gantz (W: 323-2711) about transportation to George Mason (it is accessible from the Vienna Metro Station) and further information about the course content. For questions about registration or billing call Virginia de Wolf at (W) 366-5361 or (H) 588-8812.

## Washington Academy of Sciences Meeting

The Washington Academy of Sciences will hold its September meeting in the Mary Gradon Center of American University at Massachusetts and Nebraska Avenues, NW on Thursday, September 17. Ralph Bledsoe, Special Assistant to the President and Executive Secretary of the Domestic Policy Council, will be speaking on ***Implications of Science for Domestic Policy Issues.*** There will be a reception at 6:45 p.m. followed by dinner at 7:30. The lecture will begin at 8:30 p.m. It is not necessary to attend the dinner to hear the lecture, and there is no charge for the lecture only. Call (301) 320-3621 for information and make dinner reservations at least a week in advance.

## Census Bureau's Forth Annual Research Conference Announced

The Census Bureau's Fourth Annual Research Conference (ARC IV) will be held March 20-23, 1988, at the National Clarion Hotel in Arlington, Virginia, only 1/2 mile from National Airport and three blocks from the Metro. ARC IV will comprise a mix of topics ranging from new survey techniques and census survey automation to measurement problems with foreign trade statistics. For further information contact Ms. Maxine Anderson-Brown, ARC Conference Coordinator, Office of the Director, Bureau of the Census, Washington, D.C. 20233, 301/763-1150.

## Public Data: Use It or Lose It

The International Association for Social Science Information Services and Technology (IASSIST) is a professional association of individuals who are engaged in the acquisition, processing, maintenance, and distribution of machine readable text and/or numeric social science data. Its 14th annual conference will be held in Washington, D.C. on Thursday, May 26 through Sunday, May 29, 1988. The program will consist of contributed papers, workshops and round table discussions concerned with archiving, accessing and using machine readable social science data. The conference will feature presentations on a wide variety of topics of interest to social scientists, data archivists, librarians, researchers, programmers and program administrators. The program committee is now soliciting contributions in the form of papers and proposals for panel discussions, workshops or round tables.

**SUGGESTED TOPICS**

- Data Interchange Standards
- Storage Technologies
- Public Policy Issues
- Cost of Data Access
- Microcomputer Applications
- Data Access Systems
- Networking and Electronic Messaging
- Analysis Methods
- Microsimulation
- Labor Market Surveys
- Microeconomic Panel Studies
- Administrative Data

For more information please contact: IASSIST88, c/o Pat Doyle, Mathematica Policy Research, Inc., 600 Maryland Avenue, S.W., Suite 550, Washington, D.C. 20024.

# PROGRAM ABSTRACTS

TOPIC: **Laspeyres Price Index Estimation Under an Autoregressive Model**

SPEAKER: Richard Valliant, Bureau of Labor Statistics

DISCUSSANT: Philip Kott, Department of Agriculture

DATE & TIME: Thursday, September 17, 1987; 12:30 to 2:00 p.m.

LOCATION: Room 2736, GAO Building, 441 G Street, N.W., Washington, D.C.
(Please call 523-1760 if you plan to attend in order to assure building entrance.)

SPONSOR: Methodology Section

ABSTRACT: Publication of indexes measuring changes in prices of retail, wholesale, export, and import items is an important part of many governmental statistics programs. One form of price index which is often used is the fixed base Laspeyres in which a fixed market basket of goods is priced over time. Two estimators of the Laspeyres index will be compared theoretically and empirically. One is referred to as a product estimator which employs sample data from each time period from the current period t back to some earlier base period 0. The estimator of long-term price change from 0 to t is calculated as the product of estimated 1-period price changes. The other estimator of long-term change considered here is a simpler one which uses data only from the current period's sample.

Theoretical properties are derived under a model in which long-term relative price changes for individual items have common within-stratum means and are correlated over time. The theory is tested in a simulation study in which a large number of stratified probability samples are selected from a population extracted from items priced for the U.S. Consumer Price Index. The theory and the empirical work both illustrate that there is a conflict between the goals of estimating long- and short-term price change analogous to the conflict between the estimation of level and change in other finite population problems.

---

TOPIC: **Two-Sample Linear Rank Tests for Censored Data**

SPEAKER: Gordon Lan, NHLBI

CHAIR: James Dambrosia, NINCDS

DATA & TIME: Tuesday, September 29, 1987; 1:30 to 3:30 p.m.

LOCATION: Landow Building, Conference Room E, 7910 Woodmont Avenue, Bethesda, Maryland
(Three blocks north of Bethesda Metro)

SPONSOR: Public Health and Biostatistics

ABSTRACT: A heuristic argument will be given to explain why a linear rank test is asymptotically equivalent to a generalized Mantel-Haenszel test. In many cardiac disease trials, the purpose of a new treatment is to delay the occurrence of death. For such trials, sequential monitoring of a linear rank test may lead to misinterpretation of treatment effects.

## PROGRAM ABSTRACTS (continued)

TOPIC: **Generalized Variance Functions for a Complex Sample Survey**

SPEAKER: Benjamin King, Educational Testing Service

DISCUSSANT: Steven Cohen, National Center for Health Services Research

DATE & TIME: Wednesday, October 7, 1987; 12:30 to 2:00 p.m.

LOCATION: Room 2736 GAO Building, 441 G Street, N.W., Washington, D.C.
(Please call 523-1760 if you plan to attend in order to assure building entrance.)

SPONSOR: Methodology Section

ABSTRACT: For a national survey of reading ability among young adults using a multi-stage, stratified probability sample, generalized variance functions (GVF) were estimated. That is, an attempt was made to express the estimated variance of a statistic as a function of the statistics and other characteristics of the variable of interest. With GVFs estimated from a development sample of variables, predictions of sampling variance were made for other variables in a confirmation sample and comparisons made with conventional jackknife estimates. Conclusions were drawn about the feasibility of use of GVFs, with emphasis on the margin of additional estimation error that is introduced.

## OTHER ANNOUNCEMENTS

### Fiftieth Anniversary Symposium

In 1935 a new department was created at George Washington University, the first Department of Statistics in a college of arts and sciences in the United States. On October 30, 1987, the Department of Statistics (known since 1983 as the Department of Statistics/Computer and Information Systems) will be celebrating (somewhat belatedly) with a Fiftieth Anniversary Symposium at the Marvin Center. WSS members and other statisticians are welcome to attend.

The program will begin at 9 a.m. in the Marvin Theater. Frederick Mosteller of Harvard University will deliver the keynote address. Herbert Solomon, Stanford University, will be the luncheon speaker. Other speakers will include Mitchell Gail, National Cancer Institute; Seymour Geisser, University of Minnesota; C. Terence Ireland, National Security Agency; Nancy Kirkendall, Department of Energy; Solomon Kullback, Professor Emeritus, George Washington University; Stephen Lagakos, Harvard University; Robert Shumway, University of California, Davis; Nancy Spruill, Department of Defense; and L. J. Wei, University of Michigan. The day will conclude with a wine and cheese party.

For the complete program and registration information please call Mrs. A. B. Pham, Department of Statistics/C&IS, at (202) 994-6356.

## OTHER ANNOUNCEMENTS (continued)

### NCHS Establishes Cognition and Survey Measurement Laboratory

The National Center for Health Statistics recently has made known the establishment of the National Laboratory for Collaborative Research in Cognition and Survey Measurement. Jointly funded by NCHS and the National Science Foundation, the Laboratory's mission is "to promote and advance interdisciplinary research on the cognitive aspects of survey methodology among Federal statistical agencies and the nation's universities and research centers." The Laboratory has objectives in:

*Survey Research* - to develop laboratory methods for researching the cognitive aspects of survey instruments and to improve statistical methods for measuring and controlling survey response errors;

*Cognitive Science* - to provide cognitive scientists with opportunities to collaborate in multi-disciplinary survey research projects and to use survey methods to investigate cognitive issues; and

*Federal Statistics* - to improve the quality of the nation's major statistical files used in planning, evaluating, and legislating the nation's health and social services programs.

The mission of the National Laboratory is carried out by its Collaborative Research Program, which supports university scientists in conducting problem-oriented research on cognitive issues germane to improving health and related statistics, and the Questionnaire Design Research Laboratory, which serves as a workplace in which NCHS and other Federal agencies conduct research in developing, designing, and testing their questionnaires. Further information about the Collaborative Research Program and the Questionnaire Design Research Laboratory may be obtained by writing to the Office of Research and Methodology, NCHS, Room 2-12, Center Building, 3700 East-West Highway, Hyattsville, MD 20782. (We are indebted to COPAFS for this article.)

## EMPLOYMENT COLUMN

Deadline for inserting notices is 5 (five) weeks before the publication date. Send notices and requests to: Evelyn R. Kay, 520 22nd Street, N.W., Washington, D.C. 20037, (202) 331-1153.

The Social Security Administration is seeking an individual to serve as **DIRECTOR** for its **Office of Research and Statistics** located in Washington, D.C., and Baltimore, Maryland.

The person selected for this position will be responsible for planning and directing a comprehensive policy research and analysis program and developing data systems relating to the Social Security Programs. If you are a good manager, are familiar with and sensitive to the policy development process, and have been involved in research and statistics, this may be the position for you.

The pay range for the position is $53,830 to $69,976 depending on prior pay and experience. The Social Security Administration is an equal opportunity employer. Please call Mrs. Sharon Appleby, (301) 594-3392 for additional information and/or an application.

# EMPLOYMENT COLUMN (continued)

**BIOSTATISTICIAN** -- Ph.D. or equivalent for position in Biostatistics and Data Management Section of National Cancer Institute, National Institutes of Health. Should have interest and experience/ coursework in statistical aspects of clinical trials. Experience using SAS for analysis of medical research data, and excellent oral and written communication skills are desirable. For consideration, send resume to Dr. Seth Steinberg, Building 10, Room 13C103, 9000 Rockville Pike, Bethesda, MD 20892. 301/496-9502. EOE

**MATHEMATICAL STATISTICIAN** 1529 - GS 12/13 – Position wanted in Washington, D.C. area; or private industry.

Education: Masters degree, GWU 1981.

Experience: 20 years federal government: (physical) systems reliability, acceptance sampling and statistical quality control, large-scale computerized databases, SAS/SAS graph, regression analysis (GLM), operations research. Experience in oral and written presentations.

Objective: Applied mathematical statistics. Statistical quality control/process control/ reliability.

Area of interest: Multivariate control charts.

(Please contact Evelyn Kay, code #88-1.)

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