



WSS NEWS

WASHINGTON
STATISTICAL
SOCIETY

December 2004

THE JEANNE E. GRIFFITH MENTORING AWARD

On receiving the Roger Herriot Award in June 2001, Jeanne E. Griffith said:

One of the most rewarding aspects (of Federal statistics) for me was the opportunity to promote creative activities and energies among my staff... When I have had the blessing to mentor young people in their careers, I have tried to emphasize..... (that) only they, themselves, can make the most of (the)...chances that life presents.

Dr. Griffith died in August 2001 after working for more than 25 years in the Federal statistical system. Throughout her career, and especially in her latter senior management positions at the National Center for Education Statistics and the National Science Foundation, one of Jeanne's *highest* priorities was to mentor and encourage younger staff at all levels to learn, to grow, and to recognize and seize career opportunities as they came along.

The Jeanne E. Griffith Mentoring Award has been established to encourage mentoring of younger staff in the Federal statistical system. It is presented annually, beginning in 2003, to a supervisor who is nominated by co-workers and supervisors, and chosen by the Award Selection Committee.

The award is co-sponsored by the Interagency Council on Statistical Policy, the Council for Excellence in Government, the Washington Statistical Society, the Social Statistics and Government Statistics Sections of the American Statistical Association, and the Council of Professional Associations on Federal Statistics.

Nominations for 2005 will be accepted beginning in February 2005. The last date for submission of nominations is April 1, 2005, and the Award Committee will make its determination of the award winner by May 1, 2005. The award will consist of a \$1000 honorarium and a citation, which will be presented at a ceremony arranged by the co-sponsors in June 2005.

The winning mentor will be selected for his or her efforts in supporting the work and developing the careers of younger staff. Examples of typical mentoring activities include:

Advising junior staff to help them create career opportunities, networking skills, and contacts for growth and development;

Counseling junior staff and providing resources to help develop their technical writing, analysis, presentation and organizational skills and knowledge;

Encouraging junior staff growth and career development through attendance and oral presentations at meetings with higher level officials, staffs of other agencies, professional associations, training courses, and conferences;

Motivating junior staff and building self confidence through feedback on their efforts, being a listener when that is needed, and creating a caring and supportive environment;

Serving as a role model for junior staff through professional expertise, information and insights, balancing collegial and personal roles, and including everyone across rank, race, ethnicity, and seniority.

For further information on the award, contact Ed Spar, Council of Professional Associations on Federal Statistics (COPAFS) by phone: 703-836-0404; fax: 703-684-3410; or by e-mail at copafs@aol.com. The nomination cover sheet and guidelines form—or a photocopy of it—should be attached to a nomination memorandum or letter. Forms can be obtained by contacting Ed Spar, or by downloading from the COPAFS website at <http://www.copafs.org>. All nominations should be returned to the Jeanne E. Griffith Mentoring Award Committee, c/o COPAFS, 1429 Duke Street, Alexandria, VA 22314 no later than April 1, 2005.

WSS and Other Seminars (All events are open to any interested persons)	
December 1 Wed.	Data Quality: Automated Edit/Imputation and Record Linkage
January 19 Wed.	Machine Learning Methods for Text Classification

Also available on the World Wide Wed at the following URL: <http://www.science.gmu.edu/~wss>

Announcements

JUDGES FOR THE 2005 SCIENCE FAIRS

Volunteers are needed to represent the Washington Statistical Society next spring as judges in five regional science fairs in Northern Virginia, suburban Maryland and the District of Columbia. Since 1986, WSS has provided special awards at these fairs to students whose projects demonstrate excellence in data analysis or the application of statistical methods. Those who have participated in this activity have very much enjoyed the opportunity to interact with the students and to observe the widely diverse projects which are presented. The fairs are held on a Saturday morning in mid-March to mid-April. The only time required is that one Saturday morning, plus one weekday lunchtime meeting to discuss judging strategy.

If you would like to be a science fair judge next spring, please e-mail Robert Clickner at Robertclickner@westat.com by January 25, and include your e-mail address, work and home phone numbers, your fax number and your mailing address. If you judged last spring, there is no need to contact Bob unless your e-mail address or phone number has changed. If you have any questions, please call Bob at 301-294-2815.

SIGSTAT Topics

December 8, 2004: PROC MIXED - Part 7: Generalized Linear Mixed Models (<http://www.sas.com>)

We continue with the topic begun in October 2003. In the June 2004 meeting, the difference between general linear models and models using generalized estimating equations (GEE's) was covered. The available correlation structures in PROC GENMOD were discussed and GENMOD was used to fit a longitudinal data model. To wrap up this topic, the concepts behind generalized linear mixed models are discussed and a longitudinal data model is fit using the GLIMMIX macro.

January 12, 2005 : Resampling Methods (<http://www.resampling.com/>)

Peter Bruce, developer of Resampling Stats software, will give an introduction to resampling methods and their history. He will discuss William Gossett's

original simulations that led to the development of the t-distribution and work by Fisher and Pittman (all in the early part of the last century), as well as the development of the bootstrap. The talk will include illustrations of the main resampling methods: the bootstrap and permutation procedures. This will be an overview, suited to those who are not experienced resamplers, not an in depth technical talk.

February 9, 2005 : Graphics with R (<http://www.r-project.org/>)

The ease of making graphs in R is one of its strongest attractions. Some examples of making graphs were discussed in Part I. In this second part, the topics of making graphs which are annotated with equations, of putting multiple graphs on a single page, and of using polygons will be discussed. Mike Fleming will be the speaker.

March 9, 2005: PROC UCM - Unobserved Component Models (<http://www.sas.com>)

Unobserved Component Models are very general time series models that incorporate trends, seasonality, cycles, regression effects, and autoregressive effects. Trends and seasonality can be allowed to change randomly. UCMs can be considered as generalizations of ARIMA and smoothing models. The basics of UCMs will be covered as well as some examples. PROC UCM is new in SAS/ETS version 9. Charlie Hallahan will be the speaker.

SIGSTAT is the Special Interest Group in Statistics for the **CPCUG**, the Capital PC User Group, and **WINFORMS**, the Washington Institute for Operations Research Service and Management Science. All meetings are in Room S3031, 1800 M St, NW from 12:30 to 1:30. Enter the South Tower & take the elevator to the 3rd floor to check in at the guard's desk. First-time attendees should contact Charlie Hallahan, 202-694-5051, hallahan@ers.usda.gov, and leave their name. Directions to the building & many links of statistical interest can be found at the **SIGSTAT** website, <http://www.cpcug.org/user/sigstat/>.

Note from the WSS NEWS Editor

Items for publication in the February WSS NEWS should be submitted no later than December 30, 2004. E-mail items to Michael Feil at michael.feil@usda.gov.

Program Announcement

- Topic: **Data Quality: Automated Edit/Imputation and Record Linkage (Repeated)**
- Speaker: William Winkler, US Census Bureau
- Date/Time: Wednesday, December 1, 2004/10:30 a.m. - 12:00 p.m.
- Location: Auditorium Building 3, US Census Bureau
- Abstract: Statistical agencies collect data from surveys and create data warehouses by combining data from a variety of sources. To be suitable for analytic purposes, the files must be relatively free of error. Record linkage (Fellegi and Sunter, JASA 1969) is used for identifying duplicates within a file or across a set of files. Statistical data editing and imputation (Fellegi and Holt, JASA 1976) are used for locating erroneous values of variables and filling-in for missing data. Although these powerful methods were introduced in the statistical literature, the primary means of implementing the methods have been via computer science and operations research (Winkler, Information Systems 2004). This talk provides an overview of the recent developments.

Note: This is a repeat of the September 21st WSS talk at BLS by the same title. Due to technical difficulties, remote sites were unable to participate and the speaker has graciously agreed to repeat the presentation.

Visitors will need to contact barbara.a.palumbo@census.gov at least three days before the talk to be placed on the visitors list. If driving, be sure to specify you also need to be on the parking sticker list.

Program Announcement

- Title:** Machine Learning Methods for Text Classification
- Speaker:** William Winkler, U.S. Census Bureau
- Discussant:** Daniel Gillman, Bureau of Labor Statistics
- Chair:** Alvan Zarate, Centers for Disease Control and Prevention
- Date/Time:** Wednesday, January 19, 2005 / 12:30 - 2:00 p.m.
- Location:** BLS Conference Center, Room 1. To be placed on the seminar attendance list at the Bureau of Labor Statistics you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after 'wss') by noon at least 2 days in advance of the seminar or call 202-691-7524 and leave a message. Bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station.
- Sponsor:** WSS Methodology Section
- Keywords:** Bayesian network, multinomial, training data
- Abstract:** Textual information consisting of words can be used for areas such as classification of documents into categories (e.g., industry and occupation coding), queries in web and library searches, and the record linkage of name and address lists. To use text effectively, the text might possibly be cleaned to remove typographical error and documents (records) be given a mathematical representation in a probabilistic model. This talk describes an application of Bayesian networks to classify a collection of Reuter's newspaper articles (Lewis 1992) into categories (Nigam, McCallum, Thrun, and Mitchell 2000, Winkler 2000). The results are indirectly compared with the current best-performing methods such as Support Vector Machines (Vapnik 1995, 2000) and Boosting (Freund and Schapire 1996, Friedman, Hastie, and Tibshirani 2000). For text classification, until seven years ago, the best methods in computational linguistics outperformed the best machine learning methods. Without the need to build complicated semantic or syntactical representations, the best machine learning methods now outperform the best methods in computational linguistics with a large number of widely used test decks. These make the methods much more language independent and can make them more application dependent.

JPSM Courses Announcement

The Joint Program in Survey Methodology is offering a two day course.

The title of the course is SUBJECTIVE MEASUREMENT IN SURVEYS on April 7 and 8, 2005. For 100 years, a huge amount of social science research has been done using structured questionnaires, and asking questions and interpreting answers are core activities for clinical psychologists, anthropologists, sociologists and others who do not administer structured questionnaires. During most of this time, the wording and ordering of questions has traditionally been viewed as "an art, not a science" (Hadley Cantril, 1951, p. vii). Stanley Payne (1951) cautioned that readers of his book on questionnaire design "will be disappointed if [they] expect to find here a set of definite rules or explicit directions. The art of asking questions is not likely ever to be reduced to some easy formulas (p. xi)." Thirty years later, Sudman and Bradburn (1982) agreed, saying that "no 'codified' rules for question asking exist (p. 2)."

This course will challenge that perspective and make the case that the accumulation of a great deal of knowledge throughout the social sciences about effective question-asking does indeed offer a basis for recommendations about how best to measure subjective phenomena. However, this work has been scattered across the publication outlets of numerous disciplines, and this literature has not yet been comprehensively and integratively reviewed in a central place.

The course will be engaging and useful for participants with a wide variety of backgrounds. Questionnaire designers and analysts of questionnaire data can be found in the academic world, in the business world, and in government. Individuals from all of these arenas will find this course useful for improving their work. In short, whether a researcher uses questionnaires in laboratory experiments involving 50 participants or in large-scale representative sample surveys of thousands of respondents or simply reads and interprets questionnaire-based data collected by others, this course will help him or her do better work.

The instructor is Jon Krosnick who is Professor of Communication, Political Science, and Psychology at Stanford University and director of the Stanford Methods of Analysis Program in the Social Sciences Advanced Study in the Behavioral Sciences.

Lunch and refreshments are included in the fee. Registrants will be provided with a notebook containing detailed course notes. The registration fee for employees at sponsoring agencies and affiliates is \$550, \$550 for full-time university students, and \$730 for other participants. Payment by credit card is required. Post registration payment may be done online using the registrant's confirmation number or by calling (800) 937-9320. Payment is required by March 24, 2005.

Online registration is required. Confirmation of registration will be sent after the registration form has been processed. Registration is not firm until you receive a confirmation email. The email will include directions to the course. Payment by credit card is required. Post registration payment may be done online using the registrant's confirmation number or by calling 800-937-9320. **The registration deadline is March 24, 2005.**

The course will be held at the University of Maryland Inn & Conference Center, College Adelphi, Maryland. It is located on University Boulevard at Adelphi Road, adjacent to the College Park campus of the University. Convenient parking in the UMUC garage is included in the registration fee. For hotel room reservations, call the Conference Center at (800) 727-8622.

Employment

As a service to local statisticians, WSS News provides notification of employment opportunities and description of those seeking employment here in the Washington, DC, area. Readers are encouraged to take advantage of this feature of the newsletter. The deadline for inserting notices is five (5) weeks before the publication date. Those interested should email or call Anne Peterson, at apeterson@insightpolicyresearch.com or (703) 387-3032.

Statisticians

IBM Global Services (IGS) conducts IT & Business Consulting Services (BCS), Business Transformation Services, e-business Services, Strategic Outsourcing Services and Total Systems Management Services for its various clients. Two consulting positions available for qualified candidates with degrees in statistics or related field (BS required, Master's preferred) and 4-8 years of professional experience. Candidates must also have experience using SAS and other statistical programming tools to perform data manipulation, modeling and report generation. These are client-facing roles that require excellent consulting skills, including clear communications, ability to excel in a constantly changing environment, and ability to act as a role model to motivate and work across a team. One position is for a Team Lead and requires several years experience supervising or managing 4+ staff, while transferable skills enabling the ability to lead smaller teams (2-3 people) is preferred for the second position. The positions are located in two of our IBM Northern Virginia offices. The Team Lead position requires the ability to obtain a US Security Clearance.

Interested candidates should submit their resumes via email to Suzanne McMullen, Recruiter, IBM Business Consulting Services, at suzanne.mcmullen@us.ibm.com and indicate 'SAS/Statistician Positions - Washington Statistical Society' in the email subject line.

Project Director

The National Academies seeks a project director for a study of the Social Security Representative Payee program. The position is located in Washington, DC and will last for 30 months. A committee of experts will be convened to assess the validity, reliability, practicality and appropriateness of the representative payment policies. The committee will also be asked to identify the types of representative payees who present the greatest risk of misuse of benefits, ways to reduce the risks of misuse, and ways to better protect the Social Security beneficiaries to whom representative payees have been assigned. In order to address these issues the study plan may include analyses of existing data on the representative payee program and of new data, based on a national survey, that will be gathered as a part of this study.

The project director will provide leadership for all aspects of the study, including staffing the committee of

experts and supervising other staff and contractors. Applicants for the position must have a Ph.D. in a relevant discipline, such as statistics, economics, sociology or policy analysis, or equivalent experience. The applicant must have at least 6 years of relevant experience in utilizing data from large administrative or survey data sets to answer public policy questions; experience in leading study teams; and excellent written and oral communication skills. Salary is commensurate with experience.

How to Apply:

If you are interesting in applying for the position, please e-mail or fax your resume and your contact information to Dr. Jane L. Ross at [jross@nas.edu](mailto:ross@nas.edu) or (202) 334-3829.

Contact Information:

Dr. Jane L. Ross
Director
Center for Economic, Governance, and International Studies
500 5th Street, NW
Keck Building, Room, 1125
Washington, DC 20001
Phone: 202/334-2092
Fax: 202/334-3829
E-mail: jross@nas.edu

Research Associate Positions in Cancer Surveillance

Join the staff of statisticians, epidemiologists, and demographers in the Surveillance Research Program (SRP) at the National Cancer Institute (NCI), National Institutes of Health (NIH) in the Department of Health and Human Services (DHHS) as they fulfill their mission of collecting, presenting, and interpreting the nation's cancer trends. Potential positions, full or part-time, encompass three broad areas:

- (1) Statistical Analyses – To conduct statistical analyses of pertinent data in order to answer key questions about cancer incidence, morbidity, survival, mortality, and cancer-related health status in diverse regions and in specific populations of the U.S. For example, using health survey data to obtain distributions of cancer risks using various risk models, using Surveillance, Epidemiology and End Results (SEER) data to describe emerging trends, and research in dietary assessment.
- (2) Information Technology – To assist in various IT activities, including working with several data systems related to cancer, identification of standards in software development/maintenance at federal agencies and professional organizations, and research on issues related to electronic messaging and information management.
- (3) Project management and communications --

To assist with coordination of workshops, placement of news items relevant to SRP research, and reporting statistical results. In addition, to assist the program directors in various aspects of managing the SRP portfolio of statistical grants, including administrative data entry and reporting and communication with investigators.

www.mathstat.american.edu/positions, or from the department at (202) 885-3120. American University is an Equal Employment Opportunity / Affirmative Action employer, committed to diverse faculty, staff, and student body. Women and minority candidates are strongly encouraged to apply.

Applicants must have strong quantitative skills related to the development and interpretation of health statistics. We are seeking candidates at all levels (bachelor's, master's, post-doc). Training and experience in an appropriate discipline (such as statistics, epidemiology, demography, actuarial science, behavioral science, public health research, program evaluation, or a related area) is necessary. Excellent communication and interpersonal skills are essential. Salary range is commensurate with experience, and benefits are excellent. Educational opportunities are available. The location is Rockville, Maryland. DHHS/NIH/NCI are all equal opportunity employers. Please send a cover letter, CV, brief summary of experience and interests, copies of selected publications, and three letters of reference to:

Brenda K. Edwards, Ph.D.
Surveillance Research Program, DCCPS
6116 Executive Blvd., MSC 8315
Bethesda, MD 20892-7350
 (301) 486-8506
edwardsb@mail.nih.gov

Please visit our web sites to learn more about us.
<http://surveillance.cancer.gov/>;
<http://surveillance.cancer.gov/csb/>;
<http://srab.cancer.gov/>

Tenure Track Assistant Professor

The Department of Mathematics and Statistics in the College of Arts and Sciences at American University has an opening for one tenure track assistant professor in statistics for Fall 2005.

Qualifications: earned doctorate in Statistics or in Mathematics with a specialty in Statistics by Fall 2005, as well as evidence of effective teaching and either a record of or the potential for continuing productive scholarship. Responsibilities: Teaching undergraduate and graduate level statistics courses; conducting research, advising and mentoring students, with particular sensitivity to women and minority students; institutional service.

Application review will begin November 15, 2004, and continue until the positions are filled. Submit letter of application and vitae to Search Committee, Department of Mathematics and Statistics, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016-8050. Have transcripts and three letters of reference sent directly to the department. At least one letter should specifically mention teaching experience.

All applicants are strongly advised to review full application instructions, available at



P.O. Box 752
Suitland, MD 20752

Washington
Statistical
Society

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

POSTMAN: DATED MATERIAL
Please deliver by December 1, 2004

President John Czajka	202-484-4685	Past President David Marker	301-251-4398	President-Elect Clyde Tucker	202-691-7371
	Secretary Courtney Stapleton	301-763-4142	Treasurer John Finamore	301-763-5993	
Vice Chair for District 2 of Council of Chapters Carolee Bush	202-997-2264		Council of Chapters Representative Glenn White	202-327-6414	
	Representative-at-Large Tom Mule Leslie Wallace	301-763-8322 301-738-3543	Jennifer Guarino Tancreto Michael Sinclair	301-763-4250 202-693-1153	
		WSS Program Chairs			
Agriculture and Natural Resources Mel Kollander Mike Fleming	202-537-6700 202-720-8951	Economics Linda Atkinson Karen Pence	202-694-5046 202-452-2342	Methodology Jonaki Bose Wendy Rotz	202-366-9979 202-327-7822
Public Health and Biostatistics Grant Izmirlian Jai Choi	301-496-7519 301-458-4144	Public Policy Brian A. Harris-Kojetin	202-395-7314	Quality Assurance and Physical Sciences Amrut Champaneri Eugene Burns	202-366-5998 202-366-3491
Short Courses Sylvia Dohrmann Trena Ezzati-Rice	301-610-5119 301-427-1478	Social and Demographic Statistics Judy Droitcour Kevin Deardorff	202-512-9145 301-763-6033	Statistical Computing Charlie Hallahan Dean H. Judson	202-694-5051 301-763-2057
Electronic Mail Committee Michael Greene S.V. (Vince) Massimini	301-504-7335 703- 883-5893	Data Collection Methods Sameena Salvucci	703-807-2309	Membership Fritz Scheuren John Dixon Timothy Kennel	202-320-3446 202-691-7516 301-736-6795
Webmaster Dan Jacobs	301- 403-4220	WSS NEWS Editor Michael Feil	202-690-3130	Curis Jacobs Award Amy Luo Anthony Whyde	202-327-6667 301-763-6086
Employment Anne Peterson	703-979-1191	Quantitative Literacy Carolyn Carroll	703-379-4500	Local Arrangement Kevin Cecco	202-874-0464
Video Librarian Mel Kollander	202-537-6700	Quantitative Literacy Poster Competition Archana Joshee Ryan Petska	202-327-6859 202-327-7245	Quantitative Literacy Workshop Mark Otto	301-497-5872
Social Arrangements Yves Thibaudeau	301-763-1706	WSS Committee on ASA Fellows Michael L. Cohen Fritz Scheuren Carol House Steve Cohen Dwight Brock	202-334-3765 202-320-3446 202-720-4333 301-427-1466 301-517-4026	WSS Historian & Financial Advisor Michael P. Cohen	202-366-9949
Video Coordinator John Dixon	202-691-7516			Science Fair Coordinator Bob Clickner	301-294-2815