UCLA
Statistics Newsletters

Title
Statistical Moments - Winter/Spring 2002

Permalink
https://escholarship.org/uc/item/09j8s837

Author
UCLA Department of Statistics

Publication Date
2002-12-01

Greetings from the Chair

Jan de Leeuw

This is the very first issue of the UCLA Statistics newsletter. Our department is now three years old, and it has grown tremendously over this period. I will use this occasion to give a brief overview of the various developments.

Both lower division and upper division enrollments are up by about 40%, enrollment in graduate courses and the size of the graduate program is up by more than 100%. We started in 1998 with 6 Full Time Equivalent (FTE), currently we have 7, and by the beginning of next academic year we hope to have 10 or 11. There are about ten additional teaching faculty, who participate fully in departmental affairs. Income through grants and consulting is growing.

We have created quite a few new undergraduate and graduate courses, and we hope to have our very own undergraduate major next academic year. Many of these new courses are service courses, created after consultation with client departments such as Economics and Biology (OBEE). More of these service courses, both graduate and undergraduate, are being prepared. All departments and schools on campus are invited to let us know their statistics needs.

We now have two teaching labs with iMacs, graduate students have their own desktop computers, both teaching and regular faculty have departmental computers as well. Using the labs, all lower-division courses are being reunited to 5 unit courses. The graduate curriculum and the qualifying exam structure are currently being completely overhauled.

These are all very positive developments. We had our first senate internal/external review this year, and generally the reviewers seemed to be quite happy with our performance. Our main problem is space. We continue to grow at a rapid rate, but we have no space to grow in. The department is dispersed over a number of buildings, and even within the same building people are on many different floors. Most of our space is of very low quality. This will hopefully change in two years, when Astronomy moves out of Math Sciences.

Clearly, this is an exciting time for statistics at UCLA. We hope to document the ongoing changes in subsequent issues of this newsletter. And we thank Coen, Dean, Ivo, Jose and Katie for agreeing to serve on the editorial board.

New Faculty Member Profiles

Ivo Dinov

I am an Adjunct Assistant Professor in Statistics which is a joint appointment with Neurology, UCLA School of Medicine. I have earned the following degrees: Ph.D./M.S. mathematics/statistics at Florida State University, M.S. in mathematics at Michigan Tech and a B.S. in mathematics and computer science at University of Sofia, Bulgaria. I was a predoctoral fellow in industrial engineering at FSU 1996-1996 and a postdoctoral fellow in neuroscience at UCLA 1998-2001.

I was initially trained as a mathematician, but encountered a very challenging problem of understanding the anatomy and functional connectivity of the human brain. Without statistical techniques this problem is not amenable for practical (clinical or research) purposes. I am currently doing research in statistical, mathematical and computational models for integration and analysis of biomedical data.

I am also a proud new father. My daughter, Anna-Sophia, was born in October 2001

Christopher Paul

I am presently a Lecturer in the department of statistics, teaching a variety of introductory statistics classes. I am actually a sociologist, having completed my Ph.D. in the UCLA department of sociology in June of 2001 and I consider myself an "amateur statistician" having had a fair amount of formal quantitative training during my graduate study and a greater amount of hands-on experience in my work at RAND, where I still work for a very tiny fraction of my time as a consultant. I did my undergraduate work at UCLA as well (class of '93), so I'm a real Bruin, through and through. I am currently working on publishing out of my dissertation, seeking a book contract for the entire volume, and trying to convert chapters into articles. I am also putting the finishing touches on a simulation paper assessing the various popular techniques for dealing with missing data with UCLA sociology's Bill Mason and RAND's Dan Mc-

Caffrey. The paper promises to challenge the assumed universal victory of full-Bayesian multiple imputation (a la Rubin) and provide useful and practical advice to the non-statistician data analysis practitioner.

James M. Womack

I was hired by the Department of Statistics at UCLA in the Fall of 2001 as an outside Lecturer. So far I taught Stats M11 (Fall 2001) and I'm currently assigned to teach Stats M10 this spring. I have a Ph.D. and M.A. in statistics from UC Santa Barbara, and a B.A. in mathematics from Cal State University, Northridge. I am employed at The Aerospace Corporation as an applied statistician. My areas of interest are reliability modeling, pattern classification, data visualization, and statistical intervals. I got started in statistics while in graduate school studying mathematics at UCSB. To fill some time I took a introductory course in statistics from Professor Milton Sobel. He taught a challenging course covering topics like incomplete gamma functions and group testing. I was amazed at how he would compute means and variances of data sets without the aid of computers or calculators. I continued taking courses in statistics and probability and decided to major is statistics midway through my second year in graduate school.

Hongquan Xu

I am an Assistant Professor in the Department of Statistics at UCLA. I got my Ph.D. in Statistics from University of Michigan, Ann Arbor in April 2001, and Ph.D. in Mathematics from Nankai University, Tianjin, China in July 1997. My current research interests are experimental design, computer experiments, bioinformatics and data mining. After I got my B.Sc. in pure mathematics from Nankai University in 1991, I had chance to work on some projects in pattern recognition and image processing. I found statistics is very useful; however, I did not study statistics until I was admitted to Michigan in 1997. I like playing ping pong, swimming and traveling.

Faculty News

Dick Berk & Jan de Leeuw elected fellows

Dick Berk and Jan de Leeuw have been elected Fellow of the American Statistical Association. They join Tom Ferguson, Bob Jennrich, and Don Ylvisaker. Jan de Leeuw has also been elected fellow of the IMS, joining Tom Ferguson, Bob Jennrich, Ker-Chau Li, Jim Mac-Queen, Sid Port, and Don Ylvisaker.

Staff News

Dean Dacumos chosen for Professional Development Program

In the 2001-2002 Academic Year, Dean M. Dacumos, the Department's Student Affairs Officer, is one of 26 employees chosen to participate UCLA's prestigious and competitive Professional Development Program (PDP). This Program assists employees in developing skills, knowledge and networks to improve their leadership and management skills. As a finalist for the 2001 UCLA Staff Excellence in Service Award, Dean's service is suited perfectly for this program. In addition to benefiting Dean, who recently received his Ten Years of Service pin at UCLA, this program helps the Department. The PDP Program has helped Dean implement effective methodologies in his work and introduced him to people who help the Department grow. More information on the PDP can be found on the April 2, 2002 UCLA Today (www.today.ucla.edu) newsletter. Dean is expecting his first child in 2002.

Student News

Roger Peng Wins ASA Award

Roger Peng has been working with Rick Paik Schoenberg on an NSF/EPA sponsored project titled "Fire Hazard Estimation using Point Process Methods" since the summer of 2000. The primary goal of the project is to develop and assess point process models for predicting fires in space and time. It was in 2001 that he wrote a paper with Professor Schoenberg on one small aspect of fire behavior, the renewal properties of wildfires. Many current fire hazard models assume that after a wildfire has completely burned an area, the risk of another fire at that location initially decreases and then increases linearly with time. This reasoning serves as the motivation for prescribed burning. Using fire data from the Los Angeles County Department of Public Works dating back to 1878, the two found that while the proportion of area burned does increase with the age of the fuel (i.e. the time since the last burn), there is a leveling off of that burning after a fuel age of approximately 35 years is achieved. Peng reports that one possible interpretation is that large wildfires occur when conditions are ripe, i.e. when fuel age is at least 30 to 40 years, but that there is little distinction, with regard to risk based on fuel age, between conditions that are sufficient and conditions that are extreme.

The paper is titled "Estimating the Renewal Distribution of a Spatial-Temporal Process" and it was selected as a winner of the 2002 ASA Student Paper Competition in Computing and Graphics. Roger Peng will be presenting this work at the Joint Statistical Meetings in August in New York City.

Sarah Rothenberg & Trudy Poon Awarded NSF Graduate International Outreach Award

Two M.S. students in our department, Sarah Rothenberg and Trudy Poon were recently selected to participate in a National Science Foundation (NSF) sponsored program abroad, co-sponsored by the National Institutes of Health (NIH). The summer program is designed to provide U.S. graduate students in science and engineering with first-hand research experience in Japan, Korea and Taiwan, an introduction to the science and science policy infrastructure of the respective countries, and an orientation to the culture and language. Trudy will be studying in Korea while Sarah will be in Japan. More information about the program can be found at http://www.nsf.gov/sbe/int/eap/start.htm.

Alumni News

Recent Graduates

Five individuals have recently made the jump from student to alumni. Eunice Kim, Sean Grettum, and Xiaofeng Zheng received their M.S. degrees in 2000–2001. Cheng Li and Jun Xie received their Ph.D. Expected to join them in the Spring of 2002 are M.S. students Jinrui Cui, Trudy Poon, Sarah Rothenberg, Jie Shen, Diana Wei, and Shirley Xiong and Ph.D. students Vanessa Beddo and Heidi Graziano.

Recent Publication by Amy Braverman (class of 1999)

Compressing Massive Geophysical Datasets Using Vector Quantization Journal of Computational and Graphical Statistics, 11, 2002, 44–62

Amy got her Ph.D. in Statistics in 1999, and then held a two year post-doc at JPL beginning in Novemeber 1999. In November 2001 she was was hired as a Scientist in the Earth and Space Sciences Division at JPL. Amy now works on two different instrument teams: the Multiangle Imaging Spectro Radiometer, which has been collecting data since early 2000, and the Atmospheric Infrared Sounder, which is to be launched in early May of this year. She designs algorithms to perform data reduction on the very large data sets these instruments generate. Her dissertation, "A Rate-distortion Approach to Massive Data Set Analysis" laid the groundwork for these algorithms. Amy's current research involves extending those ideas to work for even larger data sets, and better understanding their statistical properties. She collaborates with the JPL Machine Learning Systems Group, and several statisticians outside JPL on data mining and high-dimensional visualization techniques for remote sensing data. She also works routinely with applied physicists and atmospheric scientists who study climate change.

Alumni Updates

An on-line alumni update database has been developed. Please take a few minutes to update information about yourself at http://www.stat.ucla.edu/alumni/update.php. A current directory listing of alumni can be found at http://www.stat.ucla.edu/directory/alumni.php.

Recent Publications

Preprints, Papers & Reviews

For full text on these and other recent publications in our department please visit http://papers.stat.ucla.edu

"Quantitative Comparison and Analysis of Brain Image Registration Using Frequency-Adaptive Wavelet Shrinkage" by Ivo D. Dinov, Michael S. Mega, Paul M. Thompson, Roger P. Woods, DeWill L. Sumners, Elizabeth L. Sowell, and Arthur W. Toga

Abstract: In the field of template-based medical image analysis, image registration and normalization are frequently used to evaluate and interpret data in a standard template or reference atlas space. Despite the large number of image-registration (warping) techniques developed recently in the literature, only a few studies have been undertaken to numerically characterize and compare various alignment methods. In this paper, we introduce a new approach for analyzing image registration based on a selective wavelet reconstruction technique using a frequency-adaptive wavelet shrinkage. We study four polynomial-based and two higher complexity non-affine warping methods applied to groups of stereotaxic human brain structural (magnetic resonance imaging) and functional (positron emission tomography) data. Depending upon the aim of the image registration, we present several warp classification schemes. Our method uses a concise representation of the native and resliced (pre- and post-warp) data in compressed wavelet space to assess quality of registration. This technique is computationally inexpensive and utilizes the image compression, image enhancement, and denoising characteristics of the wavelet-based function representation, as well as the optimality properties of frequency-dependent wavelet shrinkage.

"Statistical Assumptions as Empirical Commitments" by Richard A. Berk and David A. Freedman

An excerpt from the introduction: Conventional statistical inferences (e.g., formulas for the standard error of the mean, t-tests, etc.) depend on the assumption of random sampling. This is not a matter of debate or opinion; it is a matter of mathematical necessity. When applied to convenience samples, the random sampling assumption is not a mere technicality or a minor revision on

the periphery; the assumption becomes an integral part of the theory. In the pages ahead, we will try to show how statistical and empirical concerns interact. The basic question will be this: what kinds of social processes are assumed by the application of conventional statistical techniques to convenience samples? Our answer will be that the assumptions are quite unrealistic. If so, probability calculations that depend on the assumptions must be viewed as unrealistic too.

"Statistical Modeling of Texture Sketch" by Ying Nian Wu, Song Chun Zhu, and Cheng-en Guo

Abstract: Recent results on sparse coding and independent component analysis suggest that human vision first represents a visual image by a linear superposition of a relatively small number of localized, elongate, oriented image bases. With this representation, the sketch of an image consists of the locations, orientations, and elongations of the image bases, and the sketch can be visually illustrated by depicting each image base by a linelet of the same length and orientation. Built on the insight of sparse and independent component analysis, we propose a two-level generative model for textures. At the bottom-level, the texture image is represented by a linear superposition of image bases. At the top-level, a Markov model is assumed for the placement of the image bases or the sketch, and the model is characterized by a set of simple geometrical feature statistics.

"High Risk and Competitive Investment Models" by F. Thomas Bruss and Thomas S. Ferguson

Summary: How should we invest capital into a sequence of investment opportunities, if, for reasons of external competition, our interest focuses on trying to invest in the very best opportunity? We introduce new models to answer such questions. Our objective is to formulate them in a way that makes results high- risk specific in order to present true alternatives to other models. At the same time we try to keep them applicable in quite some generality, also for different utility functions. Viewing high risk situations we assume that an investment on the very best opportunity yields a lucrative, possibly time-dependent, rate of return, that uninvested capital keeps its risk-free value, whereas "wrong" investments lose their value. Several models are presented, mainly for the so-called rank-based case. Optimal strategies and values are found, also for different utility functions, and several examples are explicitly solved. We also include results for the so-called full-information case, where, in addition, the quality distribution of investment opportunities is supposed to be known. In addition we present tractable models for an unknown number of opportunities in terms of Pascal arrival processes. Effort is made throughout the paper to justify assumptions in the view of applicability.

Bits && Bytes

Computing News, Information and Tips

Welcome to our first issue. This section is dedicated to disseminating computer related news from the Department of Statistics at UCLA and to provide information and tips that we hope will serve the broader readership.

Statistics WEB Resources

The web has been a principle vehicle for asserting our identity and vision, and has served to unite the department—organizing our resources under one homepage.

In 1995, under Jan de Leeuw's leadership, the department embarked on an ambitious mission of bringing statistics on-line. Our's was one of the first web sites to have statistics calculators and a statistics textbook on-line. The calculators continue to be a favorite, receiving over five thousand hits a month.

Our popular web resources are:

Seminars	http://www.stat.ucla.edu/seminars
Research	http://www.stat.ucla.edu/research
Papers	http://papers.stat.ucla.edu
Consulting	http://www.stat.ucla.edu/consult
Calculators	http://calculators.stat.ucla.edu
CIS	http://cis.stat.ucla.edu
Mailing Lists	http://lists.stat.ucla.edu

We have added the following new resources:

Support	http://www.stat.ucla.edu/support
Directory	http://www.stat.ucla.edu/directory
Calendars	http://www.stat.ucla.edu/calendars
Forums	http://www.stat.ucla.edu/forums

Resources being developed or planned for the future include: Preprints database, Cluster tutorials, Expanded on-line labs.

If you have suggestions or ideas, or, are a member of the Department and have a resource you would like to contribute, please e-mail us at webstaff@stat.ucla.edu.

New resources will be announced here as they become available.

The Statistics Learning Resource Center Opens

The Learning Resource Center opened this Spring 2002. A relaxed and interactive learning environment, with assistants and reference books on hand, the lab is well suited for working on Statistics labs, holding tutoring sessions and office hours. The schedules for the Learning Resource Center and the Instructional Computing Lab are at http://www.stat.ucla.edu/esources.

Statistics Labs Featured in Consortium's Mac OS X Web Site

Our Statistics labs have been running Mac OS X since the Winter 2002 quarter. The labs and their deployment are featured in the Higher Education Mac OS X Lab Deployment Initiative web site hosted by the University of Utah. Find out how we and twenty-three participating universities have deployed our Mac OS X labs at http://www.macosxlabs.org.

Statistics Cluster

The Statistics cluster currently includes one file server (G4 933 Mhz, 1280 MB) and 16 computational nodes (G4 400 MHZ, 1280 MB).

Software includes:

PVM, R (including RPVM), LAM/MPI, Maui Scheduler Molokini Edition.

Scientific libraries include:

ATLAS, BLAS, LAPACK, BLACS, SCALAPACK and SPRNG.

Compilers include: G77 and GCC.

Spec marks can be seen at:

http://exodus.physics.ucla.edu/ appleseed/benchmarks.html

Cluster projects (past and present) include:

- *Parallel Computing in Statistics*, Vanessa Beddo
- *Multivariate Analysis with Large Datasets*, Coen Bernaards, Vanessa Beddo
- *Liquid Association–New Bioinfomatics Tool*, Ker-Chau Li, Shin-Sheng Yuan
- *Search for Optimal Designs*, Hongquan Xu
- *Parallel Particle In Cell Modeling of Semiclassical Quantum Models*, Dauger

Help files and documentation for the cluster are under development and will be announced here when they become available. To schedule time or obtain more information about the cluster contact: cluster@stat.ucla.edu. The cluster's home page is: http://www.stat.ucla.edu/cluster.

The Tip Off: Creating PDF Files in Mac OS X

A built in feature of Mac OS X is the ability to see documents in Portable Document Format (PDF). In fact, everything you see in Mac OS X *is* PDF. The added benefit is we can also save a document as PDF fairly easily using the print dialog.

How To Save a Document as PDF
1. Open the document
2. Select **Print...** from the **File** menu
3. Choose **Output Options** from the options menu
4. Click on the **Save as File** check box to mark it
5. Choose **PDF** from the **Format:** menu
6. Click the **Save...** button
7. A **Save to File** window will appear. Type in the name of the file in the **Save As:** field and navigate where to save the file with the **Where:** pop-up selector.
8. Click on the **Save** button.

That's it! A PDF file has been saved on your hard disk with the name and location you chose. Typically double-clicking on the PDF document will open it and display it. Two applications exist to open PDF files: Mac OS X's built-in viewer **Preview** and Adobe's **Acrobat**, but more about them later.

Calendar

Events in the Department

ASA's Joint Statistical Meeting 2002

We would like to announce that our UCLA Department of Statistics faculty and students will be presenting and chairing at the American Statistical Association's JSM (Joint Statistical Meetings) 2002 in New York City from August 11th to August 15h! Their schedules are as follows:

FACULTY:

Coen Bernaards:
> Session 255, August 11th @ 3:35pm
> Presenting: "Imputation in Surveys"
> Session 355, August 14th @ 3:50pm
> Chairing: "Interval Estimation"

Robert Gould:
> Session 28, August 11th @ 4:00pm
> Presenting: "Initiatives Supporting the Teaching of Statistics"

Bengt Muthen:
> Session 9, August 11, 2002 @ 2:00pm
> Presenting: "Applied Statistics Outside the Mainstream"

Rick Paik Schoenberg:
> Session 70, August 12th @ 8:30am
> Presenting: "Environmental Applications"

Hongquan Xu:
> Session 180, August 13th @ 10:05am
> Presenting: "Environmental Design"

STUDENTS:

Vanessa Beddo:
> Session 52, August 11th @ 4:00pm
> Presenting: "Data Mining and Estimation"

Ray-Bing Chen:
> Session 196, August 13th @ 10:30am
> Presenting: "Data Visualization, Data Mining and Computational Statistics"

Stephen Erickson:
> Session 282, August 14th @ 8:30am
> Presenting: "Bayes Analysis of Microarray Data"

Roger Peng:
> Session 200, August 13th @ 10:30am
> Presenting: "Student Paper Competition Winners: Statistical Computing and Statistical Graphics"

Shin-Sheng Yuan:
> Session 257, August 13th @ 3:20pm
> Presenting: "Understanding Microarray Data"

Congratulations and we wish you the best!

Important Upcoming Dates

First Day of Spring Classes	April 1, 2002
Memorial Day Holiday	May 27, 2002
Dept Qualifying Exams	June 1, 2002
Last day to file theses and dissertations for Spring	June 3, 2002
Spring Finals Week	June 10-14, 2002
Graduation (Hooding Ceremony)	June 15, 2002
Independence Day Holiday	July 4, 2002
Summer Session A (8 week)	June 24 - August 16, 2002
Summer Session C (6 week)	August 5 - September 13, 2002
First Day of Fall Classes	Sept 26, 2002

UNIVERSITY OF CALIFORNIA, LOS ANGELES
DEPARTMENT OF STATISTICS
8130 MATH SCIENCES BUILDING
BOX 951554
LOS ANGELES, CALIFORNIA 90095-1554

MA-97