THE ISBA BULLETIN

Vol. 8 No. 2 *June 2001*

The official bulletin of the International Society for Bayesian Analysis

A WORD FROM THE PRESIDENT

by Alicia Carriquiry
ISBA President
alicia@iastate.edu

Dear ISBA member:

Greetings! I hope your summer (or winter!) is going well.

Today I would like to point you to two ISBA matters that require your attention and that are highlighted in the Bulletin.

First, the ISBA Board of Directors has assembled a Nominations Committee that will propose candidates for election to various ISBA posts. The Nominations Committee is chaired by Phil Dawid, ISBA Past President; you will find the list of committee members in an article later in the Bulletin. The positions that need to be filled come January 1, 2002, are President Elect, Treasurer, and four Directors.

In addition, you will also find an announcement for the 7th Valencia (V7) meeting, which in 2002 will be co-sponsored by ISBA. A new agreement between ISBA and the Valencia International Conference Organizing Committee has resulted in a significant change in the structure of the V7 scientific program, and I urge you to read all about that in the article that follows. ISBA will not only be an "official" co-sponsor of the meeting, but will also be in charge of selecting and organizing a major portion of the presentations. Hal Stern has agreed to lead this effort on behalf of ISBA, and we look forward to a week of productive scientific endeavors in a spectacular venue.

Your input on the two matters above is, as always, greatly appreciated and valued. The Nominations Committee is urgently soliciting suggestions for nominees for the various ISBA elective offices; I hope that you will get in touch with Phil with your suggestions. Also please check the ISBA web site at www.bayesian.org for the latest news regarding the Valencia meeting, and think about participating by submitting an abstract.

I am personally always delighted to hear from the ISBA members, so do not hesitate to drop me a line with comments, suggestions or concerns.

All the best.

A WORD FROM THE EDITOR

by Fabrizio Ruggeri
ISBA Bulletin Editor
fabrizio@iami.mi.cnr.it

After discussing in past issues the birth of a new Bayesian era, now we proudly announce a new, Bayesian, geography: the next Valencia meeting will be in the Canary Islands. Jose Bernardo will be always remembered for it (besides reference priors, of course!).

Contents

ISBA NOMINATING COMMITTEE 2001

This year's Nominating Committee has recently been appointed by the Board. Its task is to put forward to the membership two nominees for each vacancy that will arise as from 1 January 2002, namely: President-Elect, Treasurer (to replace Valen Johnson), and four members of the Board of Directors (to replace Mark Berliner, Petros Dellaportas, Jayanta Ghosh and Sylvia Richardson). The procedures of the Committee are governed by Section D of the ISBA By-laws, which are displayed on the ISBA website at `www.bayesian.org`

The composition of the Nominating Committee is as follows: Philip Dawid (Chair, University College London, UK), Gabriel Huerta (CIMAT, Mexico), Katja Ickstadt (University of Darmstadt, Germany), Rosangela Loschi (Universidade Federal de Minas Gerais, Brazil), Kerrie Mengersen (University of Newcastle, Australia), Lawrence Pettit (Goldsmiths College, UK) and Hal Stern (Iowa State University, USA).

If any member of ISBA wishes to make any suggestions as to suitable nominees, please do so, with a supporting statement, by e-mail to Philip Dawid (`dawid@stats.ucl.ac.uk`) as soon as possible.

After the Nominating Committee has drawn up its list, there will be an opportunity for members to petition to add further candidates.

7TH VALENCIA INTERNATIONAL MEETING ON BAYESIAN STATISTICS

2002 ISBA INTERNATIONAL MEETING

Spain, June 2 - 6, 2002

The Valencia 7 meeting will be held on the beautiful island of Tenerife, in the Spanish Canary Islands, during June 2 - 6, 2002. It will be jointly sponsored by the University of Valencia and ISBA, and organized by the Valencia International Conference Organizing Committee.

In addition to the usual invited presentations and the ever popular evening contributed poster sessions, the scientific program of V7 will include, for the first time, Selected Contributed Papers (SCPs), to be presented orally. A SCP Selection Committee, appointed by ISBA and chaired by Hal Stern, will be in charge of reviewing the abstracts submitted, and selecting about 50 for presentation. Detailed instructions for submitting abstracts will shortly be posted on the Valencia 7 web site at `www.uv.es/valencia7` and its US mirror site at `www.stat.duke.edu/valencia7`, and also on the ISBA web site at `www.bayesian.org`.

The deadline for submission of abstracts for consideration by the SCP Selection Committee is October 15, 2001. Authors of selected contributed papers will be notified during the second half of January 2002. Unselected contributed papers will be eligible for poster presentation. The deadline for poster submissions is May 1, 2002. For both SCP and poster submissions, at least one author must have registered for the meeting by the relevant deadline.

All SCPs and posters presented at the meeting can be considered for publication in the Valencia 7 Proceedings volume, subject to a stringent review process. The Lindley Prize will be awarded for the best contributed paper or poster published in the Proceedings; for further details see `www.bayesian.org/awards/LindleyPrize.html`.

BULLETIN BY E-MAIL

SENDING A MESSAGE TO `isba@iami.mi.cnr.it`, SPECIFYING THE PREFERRED FORMAT (POSTSCRIPT, COMPRESSED POSTSCRIPT OR PDF)

LAGUNA BEACH

by Dale J. Poirier
Local Organizer
poirier@uci.edu

A Regional ISBA Meeting was held April 5-8, 2001 in Laguna Beach, California - a picturesque ocean side town that began as an artist colony a century ago. The setting was the historic Hotel Laguna on Main Beach. Local sponsors consisted of the Institute for Mathematical Behavioral Sciences, the Department of Economics, the Department of Logic and Philosophy of Science, and the School of Social Sciences, all of the University of California, Irvine. Generous financial support by the Section on Bayesian Statistical Science of the American Statistical Association was also received. The focus of the meeting was on Bayesian applications in the behavioral and social sciences, especially cognitive science, economics, and philosophy of science. There were 70 attendees. With the help of Bill Batchelder and Brian Skyrms I was able to organize eight plenary sessions each consisting of two papers and two discussants. There were also twenty poster presentations. The complete program can be found at www.socsci.uci.edu/bayesian. The capable administrative work of Janet Phelps guaranteed a smooth running operation. This meeting had as its goal to reach out and attract new members to ISBA, particularly from areas that were under represented in ISBA. In this regard it was a great success - substantially more than half of the attendees were new members to ISBA. In the usual ISBA tradition, the conference closed with a banquet followed by an after dinner talk by Arnold Zellner and an impromptu singing of Bayesian songs led by Brad Carlin at the keyboard.

VARENNA

by Susie Bayarri
Scientific Committee Member
susie.bayarri@uv.es

The *Second Workshop on Bayesian Inference in Stochastic Processes*, endorsed by ISBA, took place in Varenna (Italy), May 31-June 2, 2001. It was a great success, with nearly 90 participants from Germany, Greece, Ireland, Italy, Mexico, The Netherlands, Poland, Spain, UK, USA, and Venezuela.

Talks and posters dealt with i) inferences in stochastic process models, ii) inferences using stochastic processes as priors, and iii) methods and results for Bayesian inference requiring stochastic processes in their development. Well established areas like time series, spatial and/or temporal models, particle filtering, non-parametrics, MCMC methods, etc. were well represented. There was also ample room for areas less frequently seen like queueing models, diffusion and birth/death processes, Markov and semi-Markov models, martingale and point processes, etc. as well as many novel applications of other stochastic processes. Most papers had a strong applied motivation, played a fundamental theoretical role or were potentially important in crucial areas in today's world.

Varenna is a beautiful, picturesque little village hanging over Lake of Como, not so spoilt by tourism (nor Bayesians!) yet. The venue of the meeting, "Villa Monastero", was a Cistercian convent founded in 1208. In 1645 the old convent was changed into a splendid abode, which kept being improved over the years yielding the beautiful building that is today. The Villa stands on the shore of the lake in a magnificent park that stretches over more than a kilometer. In spite of the glorious surroundings, the talks were well attended, even after the great banquets at night and the casual and delicious lunches on the cafes and trattorias along the shore of Varenna. The organizers also treated us with a lovely boat trip to Bellagio, and along the lake to see the spectacular Villas, and ending with a great dinner in a romantic, old hotel in Menaggio. The nearly full moon reflecting on the flat lake made for a memorable turning home, even though Bayesians and boats are known not to mix well!

The Scientific Committee (Bayarri, Consonni, Higdon, Muliere, Petrone, Regazzini, Rios Insua, Ruggeri, West and Wiper) thanks all the attendants for great presentations, fruitful exchanges and a great atmosphere, both scientifically and socially. Most of us, however, had little to do, as most of the meeting was splendidly and efficiently organized by Fabrizio Ruggeri and Sonia Petrone, who did virtually all the work. To both of them, and to the local "kids" (Bottolo, Mercaldi, Palomo and Varini) who so nicely helped, our deepest thanks.

See you all at the Third Workshop!

DAVID RIOS INSUA

by Mike Wiper
mwiper@est-econ.uc3m.es

This time we are interviewing David Rios Insua, who is Professor of Stats and Operations Research and Head of Department at Rey Juan Carlos University (URJC). The interview took place at various bars in Madrid, Varenna and over the net. This interview was meant to appear in our last issue as interviewers, but time (and Fabrizio's) pressure requires it now.

1. Your father is one of the founding fathers of Stats in Spain, your older brother is Head of the OR group at Madrid Technical University, a sister is a statistician,... Why did you want to become a statistician? Is it an advantage or a disadvantage to have a famous father in the same profession?

To start with, I should say that my father was the introducer of modern Stats in Spain, some time in the late 40's, somehow by chance, as he wanted to be a professor in Madrid (at that time he was into functional analysis in some provincial university in the, by then, isolated Spain). I am not sure that I may qualify as a statistician; my first degree was in Operations Research; my PhD was in Decision Analysis,... later with time I had to learn a bit of Stats. As for the advantage or not, my father retired when I was a freshman and ever since finishing my MSc at where he used to taught, I think I've been about ten times, and that was fourteen years ago... It has had obvious advantages: a great library at home, connections with other scientists abroad, scientific and professional discussions from a broader perspective ... but it had disadvantages: some people would doubt when I was doing Ok in exams, I had to be specially careful with professors not particularly friendly with my dad...

2. I guess your father is not Bayesian. Why did you turn to the dark side?

Well, I guess he's now a Bayesian, after my brainwashing... He's actually been always very open to new ideas. For example, the great founding fathers, like Savage, Raiffa or De Finetti, visited Spain, through his intervention; several reputed Bayesians did their (Bayesian) theses under his supervision; his paper at Bayesian Stats 1 on the foundations of robust Bayes is well-known; he's been always interested in Decision Theory,... As for myself, the University I did my undergrad was extremely traditional and I just had a couple of weeks of Bayesian Stats over five years. I remember I had problems in understanding confidence intervals... My interest came through the backdoor of Decision Theory under Uncertainty, my favourite topic for a long while; the final push came through Simon French, my PhD advisor. I remember my third or fourth day at Manchester, when I mentioned him that I could investigate some non-Bayesian decision making methods and he told me: "You may do it, but if you keep on doing this you'll have to fly back to Madrid in a few days". I guess that's a good reason for becoming a Bayesian. Later on the practice of Decision Analysis and Stats has convinced me of the relevance of Bayesian ideas in applied work.

3. Apart from family influences are there any other important people who have been important in your career? I guess you have done quite a bit of collaborative research?

I've already mentioned Simon, my maestro. I guess I've been fortunate enough to work with some twenty collaborators and all of them have helped in shaping my thinking. At Purdue, Duke and Milano, I had the opportunity to interact with a lot of people who taught me a lot as well.

4. You have worked in a large variety of fields such as decision theory, neural nets, robustness, time series applications... Maybe you could tell us a bit about what interested you in each of these areas. Do you find you can use your experience in one area to help with your work in another? Are you more of a decision analyst than a statistician?

As I mentioned I am unsure of being a statistician, nor a decision analyst. I like to picture myself as Mr. Wolf in Tarantino's Pulp Fiction. 'I am Mr. Wolf and I solve problems'. I like solving real problems and

in such sometimes I need to forecast, sometimes I need to simulate, sometimes optimise, sometimes do differential equations, most of the times a bit of everything. I guess this interest comes from my days at IIASA. Before then I was doing fairly theoretical work on the foundations of Decision Theory; there I had the opportunity to work on a large international river management project and I learned the fun of doing applied stuff: you do useful things for other people and, if the problem is novel enough, you may innovate methodologically.

5. You have set up many joint projects with industry, for example in reservoir management, customer satisfaction in banks,... Is it easy to get your statistical ideas across to professionals and do you find you have to explain things very differently to bankers, engineers,...?

When we started consulting we had the idea of explaining all kinds of technicalities to managers. That's quite wrong, as they will very likely won't have time to understand what's going on in full detail. They just want to see results in an understandable manner. I have found that I am not particularly good at that job, so now one of my colleagues does the talk show and I work more at the backoffice. For me, what is important is that we have a general framework to solve problems.

6. As head of a big department (150 people) and of various research projects, I guess that you must spend a lot of time on administration. What do you prefer: administration, research or teaching?

I enjoy any activity in which I may learn and it's not stressful. After four years administrating, the learning curve has declined and stress remains... I guess I'll give up administrating pretty soon. I still learn a lot by teaching and find little stress in teaching. Finally, no stress and so many new things to be done in research ....

7. In the programmes that you run at URJC, you try only to teach Bayesian Stats and Decision Analysis. What are the advantages of this approach when compared to the traditional one of teaching classical Stats with may be one Bayesian approach? Do you see any problems with your approach?

When I moved to URJC School of Engineering, one of the good things was that it was a brand new University with possibilities of innovation in many ways, including teaching. Therefore I promoted the use of Bayesian methods in all engineering statistical courses. The experience is going pretty well and I feel students learn to reason probabilistically and statistically and apply critically the methods in their areas of interest, with little difficulty. They understand subjective probability and utility well and enjoy using simulation methods. Our standard course is at a level intermediate between Don Berry's introductory course and Gelman et al's, book with the aid of R. There are several problems with the approach. First, there's a lack of an appropriate textbook for our level of students (a bit more sophisticated mathematically than Berry's intended audience, but not as much as Gelman's), specially in Spanish, but we are gradually developing notes; in some areas, e.g. chemical engineering we still lack nice examples, but we are on it; perhaps, what it worries us more is that when students go to their jobs they will face a frequentist world... but we teach them some of the classical stuff based on noninformative priors.

8. You are still a young guy. What about your plans and ambitions for the future?

Resigning as Head. I have a couple of good ideas for Bayesian .com companies. I am interested now in Electronic Democracy, as a way to distribute rational decision making across the net. I really look forward to launch our TED project (Towards Electronic Democracy: Internet based complex support). I would also like to cooperate to push up Statistical teaching and the statistical profession in Spain, which I think is a bit on the downside. Also we still need some time to consolidate our Bayesian group at URJC (the only fully Bayesian group in the country), and then probably would be time to move to another place. At some point, I would like to consult for the European Government... besides improving my drumming and salsa dancing.

`9. And what about the future of Bayesian Statistics and Decision Theory in general?`

Very bright!!! But we should still try to get more coverage in the media (as recently in The Economist, NYT and others) and develop more Bayesian technology. I have heard of fuzzy cameras, washing machines, tube controllers and toasters, but not of Bayesian alikes. Also, there are many important areas which we need to address, for example official statistics, and interact more with IT people in areas like data mining, web engineering,...

`10. And finally, what about the future of ISBA? Would you like to see a Bayesian journal?`

I feel ISBA should address more the media making publicity of the excellent applied and applicable work that our members are doing. I think there are too many journals already around.

BAYESIAN APPLICATIONS IN *JASA*, 1996-2000

by Kate Cowles
`kcowles@stat.uiowa.edu`

➤ Invitation from the Associate Editor

The Applications section of this Bulletin has multiple purposes. One aim, of course, is to showcase the use of Bayesian methods to address substantive problems in diverse applications areas. At least as important is to share aspects of our applied Bayesian work that will be interesting, instructive, and encouraging to fellow practitioners. As associate editor for the Applications section, I invite ISBA members to contribute short articles that address such questions as:

1. What were the major challenges involved in carrying out your Bayesian analysis?

2. On the road to developing a workable solution, did you try modeling approaches, computational strategies, etc. that proved infeasible? What lessons did you learn?

3. If your applications area is one in which Bayesian methods are not yet widely used and accepted, how did you communicate successfully with subject-area experts, including journal referees?

Please contact me at the e-mail address above if you have an idea for a contribution.

➤ Bayesian Applications in *JASA*

In this issue, we will look briefly at Bayesian applied work appearing recently in a statistical forum — specifically in the Applications and Case Studies section of the *Journal of the American Statistical Association* during the last 5 years. Of the 131 papers published in Applications and Case Studies in the March, 1996, through Dec., 2000, issues of JASA, 48 (37%) were Bayesian.

Almost half (22) of the Bayesian papers concerned medical topics, with pharmacokinetic and pharmacodynamic modeling (5 papers), medical imaging (4 papers), health care policy and management (3 papers), and disease mapping (2 papers) being the most frequent medical subclassifications. Economics (5 papers) and wildlife and fish population modeling (4 papers) appeared often. Subject areas represented by two papers each were demography, ecology, educational and psychological testing, the law, marketing, sports, and surveys.

Two papers must be mentioned individually. West, Prado, and Krystal (1999) won the 1999 American Statistical Association Award for Outstanding Statistical Application. (Due to incorrect reproduction of several figures in the June, 1999, issue, the entire paper was reprinted in the December, 1999, issue. It is counted only once in the above tallies.) With near prescience, Gelman, King, and Boscardin (1998) considered the probability of a tie in a state's vote in a presidential election when that state would determine the outcome in the Electoral College.

The number, variety, and quality of the Bayesian papers in JASA Applications and Case Studies over the past five years are gratifying. However, Bayesians are working in many substantive areas not represented, including humanities, social sciences, and physical sciences. Let us aim over the next 5 years for at least 50% of Applications and Case Studies papers to be Bayesian with an even broader range of subject areas addressed.

References are provided in chronological order.

➤ **References**

KAHN, M.J. and RAFTERY, A.E. (1996) Discharge Rates of Medicare Stroke Patients to Skilled Nursing Facilities: Bayesian Logistic Regression with Unobserved Heterogeneity. *JASA*, **91**:29-41.

JOHNSON, V.E. (1996) On Bayesian Analysis of Multirater ordinal Data: An Application to Automated Essay Grading. *JASA*, **91**:42-51.

WAKEFIELD, J. (1996) The Bayesian Analysis of Population Pharmacokinetic Models.*JASA*, **91**:62-75.

COWLES, M.K., CARLIN, B.P. and CONNETT, J.E. (1996) Bayesian Tobit modeling of Longitudinal Ordinal Clinical Trial Compliance Data. *JASA*, **91**:86-98.

WAKEFIELD, J. and BENNETT, J. (1996) The Bayesian Modeling of Covariates for Population Pharmacokinetic Models. *JASA*, **91**:917-927.

QIAN, W., TITTERINGTON, D.M. and CHAPMAN, J.N. (1996) An Image Analysis Problem in Electron Microscopy. *JASA*, **91**:944-952.

PENG, F., JACOBS, R.A. and TANNER, M.A. (1996) Bayesian Inference in Mixtures-of-Experts and Hierarchical Mixtures-of-Experts Models with an Application to Speech Recognition. *JASA*, **91**:953-960.

GELMAN, A., BOIS, F. and JIANG, J. (1996) Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions. *JASA*, **91**:1400 - 1412.

GHOSH, M., NANGIA, N. and KIM, D.H. (1996) Estimation of Median Income of Four-Person Families: A Bayesian Time Series Approach. *JASA*, **91**:1423-1431.

NANDRAM, B., SEDRANSK, J. and SMITH, S.J. (1997) Order-Restricted Bayesian Estimation of the Age Composition of a Population of Atlantic Cod. *JASA*, **92**: 333-40.

WEIR, I.S. (1997) Fully Bayesian Reconstructions from Single-Photon Emission Computed Tomography Data. *JASA*, **92**:49-60.

HSIAO, C. and TAHMISCIOGLU, A. K. (1997) A Panel Analysis of Liquidity Constraints and Firm Investment. *JASA*, **92**:455-465.

NORMAND, S-L T., GLICKMAN, M.E. and GATSONIS, C. (1997) Statistical Methods for Profiling Providers of Medical Care: Issues and Applications. *JASA*, **92**:803-814.

MALEC, D., SEDRANSK, J., MORIARITY, C.L. and LECLERE, F.B. (1997) Small Area Inference for Binary Variables in the National Health Interview Survey. *JASA*, **92**: 815-826.

GELFAND, A.E., MALLICK, B.K. and POLASEK, W. (1997) Broken Biological Size Relationships: A Truncated Semiparametric Regression Approach with Measurement Error. *JASA*, **92**:836-845.

DOMINITZ, J. and MANSKI, C.F. (1997) Using Expectations Data to Study Subjective Income Expectations. *JASA*, **92**:855-867.

DAPONTE, B.O., KADANE, J.B. and WOLFSON, L.J. (1997) Bayesian Demography: Projecting the Iraqi Kurdish Population, 1977-1990. *JASA*, **92**: 1256-1267.

MULLER, P. and ROSNER, G.L. (1997) A Bayesian Population Model with Hierarchical Mixture Priors Applied to Blood Count Data. *JASA*, **92**:1279-1292.

GELMAN, A., KING, G. and BOSCARDIN, W.J. (1998) Estimating the Probability of Events that have Never Occurred: When is Your Vote Decisive? *JASA*, **93**: 1-9.

GLICKMAN, M.E. and STERN, H.S. (1998) A State-Space Model for National Football League Scores. *JASA*, **93**:25-35.

FAUCETT, C.L., SCHENKER, N. and ELASHOFF, R.M. (1998) Analysis of Censored Survival Data with Intermittently Observed Time-Dependent Binary Covariates. *JASA*, **93**:427-437.

RAFTERY, A.E. and ZEH, J.E. (1998) Estimating Bowhead Whale Population Size and Rate of Increase from the 1993 Census. *JASA*, **93**:451-463.

STASNY, E.A., KADANE, J.B. and FRITSCH, K.S. (1998) On the Fairness of Death-Penalty Jurors: A Comparison of Bayesian Models with Different Levels of Hierarchy and Various Missing-Data Mechanisms. *JASA*, **93**:464-477.

GELMAN, A., KING, G. and LIU, C. (1998) Not Asked and Not Answered: Multiple Imputation for Multiple Surveys. *JASA*, **93**: 846-857.

BRADLOW, E.T., WEISS, R.E. and CHO, M. (1998) Bayesian Identification of Outliers in Computerized Adaptive Tests. *JASA*, **93**: 910-919.

COHEN, J., NAGIN, D., WALLSTROM, G. and WASSERMAN, L. (1998) Hierarchical Bayesian Analysis of Arrest Rates. *JASA*, **93**:1253-1259.

IBRAHIM, J.G., RYAN, L.M. and CHEN, M-H. (1998) Using Historical Controls to Adjust for Covariates in Trend Tests for Binary Data. *JASA*, **93**:1282-1293.

LIU, J.S., NEUWALK, A.F. and LAWRENCE, C.E. (1999) Markovian Structures in Biological Sequence Alignments. *JASA*, **94**: 1-15.

DOMINICI, F., PARMIGIANI, G., WOLPERT, R.L. and HASSELBLAD, V. (1999) Meta-Analysis of Migraine Headache Treatments Combining Information From Heterogeneous Designs. *JASA*, **94**:16-28.

DANIELS, M.J. and GATSONIS, C. (1999) Hierarchical Generalized Linear Models in the Analysis of Variations in Health Care Utilization. *JASA*, **94**:29-42.

BRADLOW, E.T. and ZASLAVSKY, A.M. (1999) A Hierarchical Latent Variable Model for Ordinal Data from a Customer Satisfaction Survey with "No Answer" Responses. *JASA*, **94**:43-52.

ALLENBY, G.M., LEONE, R.P. and JEN, L. (1999) A Dynamic Model of Purchase Timing with Application to Direct Marketing. *JASA*, **94**: 365-374.

WEST, M., PRADO, R. and KRYSTAL, A.D. (1999) Evaluation and Comparison of EEG Traces: Latent Structure in Nonstationary Time Series. *JASA*, **94**: 375-387 (reprinted in **94**:1083-1095).

BERRY, S.M., REESE, C.S. and LARKEY, P.D. (1999) Bridging Different Eras in Sports. *JASA*, **94**:661-676.

GUO, W., WANG, Y. and BROWN, M.B. (1999) A Signal Extraction Approach to Modeling Hormone Time Series with Pulses and a Changing Baseline. *JASA*, **94**: 746-756.

GIVENS, G.H. (1999) Multicriterion Decision Merging: Competitive Development of an Aboriginal Whaling Management Procedure. *JASA*, **94**:1003-1014.

BOATWRIGHT, P., MCCULLOCH, R. and ROSSI, P. (1999) Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model. *JASA*, **94**:1063-1073.

DATTA, G.S., LAHIRI, P., MAITI, T. and LU, K.L. (1999) Hierarchical Bayes Estimation of Unemployment Rates for the States of the U.S. *JASA*, **94**:1074-1082.

IVERSEN, D.S., PARMIGIANI, G., BERRY, D.A. and SCHILDKRAUT, J. (2000) Genetic Susceptibility and Survival: Application to Breast Cancer. *JASA* **95**:28-42.

ELLIOTT, M.R. and LITTLE, R.J.A. (2000) A Bayesian Approach to Combining Information from a Census, a Coverage Measurement Survey, and Demographic Analysis. *JASA*, **95**:351-362.

TRENKEL, V.M., ELSON, D.A. and BUCKLANT, S.T. (2000) Fitting Population Dynamics Models to Count and Cull Data using Sequential Importance Sampling. *JASA*, **95**:363-374.

GENOVESE, C. (2000) A Bayesian Time-Course Model for Functional Magnetic Resonance Imaging Data (with discussion). *JASA*, **95**:691-719.

AURENEN, K., ARJAS, E., LEINO, T. and TAKALA, A.K. (2000) Transmission of Pneumococcal Carriage in Families: A Latent Markov Process Model for Binary Longitudinal Data. *JASA*, **95**:1044-1053.

DUNSON, D.B. and ZHOU, H. (2000) A Bayesian Model for Fecundability and Sterility. *JASA*, **95**:1054-1062.

BEST, N.G., ICKSTADT, K. and WOLPERT, R.L. (2000) Spatial Poisson Regression for Health and Exposure Data Measured at Disparate Resolutions. *JASA*, **95**:1110-1118.

SANSÓ, B. and GUENNI, L. (2000) A Nonstationary Multisite Model for Rainfall. *JASA*, **95**:1089-1100.

NANDRAM, B., SEDRANSK, J. and PICKLE, L.W. (2000) Bayesian Analysis and Mapping of Mortality Rates for Chronic Obstructive Pulmonary Disease. *JASA*, **95**:1110-1118.

KVAM, P.H., TIWARI, R.C. and ZALKIKAR, J.N. (2000) Nonparametric Bayes Estimation of Contamination Levels Using Observations from the Residual Distribution. *JASA*, **95**:1119-1126.

BAYESIAN ROBUSTNESS

by James O. Berger, David Ríos Insua and Fabrizio Ruggeri

berger@stat.duke.edu
d.rios@escet.urjc.es
fabrizio@iami.mi.cnr.it

Robust Bayesian Analysis [1] is concerned with the sensitivity of the results of a Bayesian analysis to the inputs for the analysis. The early 90's was the "Golden Age" of robust Bayesian analysis, in that many statisticians were highly active in research in the area, and rapid progress was being achieved. It was a major topic in many meetings, even some non-Bayesian meetings, and there were three meetings explicitly dedicated to the topic: the Workshop on Bayesian Robustness, held at Purdue University in 1989, and two International Workshops on Bayesian Robustness, held in Italy (Milano, 1992, and Rimini, 1995). The proceedings of the latter two conferences were published, respectively, in the *Journal of Statistical Planning and Inference* (vol. 40, no. 2& 3), and in Berger *et al.* (1996), as an IMS Lecture Notes volume. An issue of the journal *Test* (vol. 3, 1994) was mainly devoted to a review paper by Berger on robust Bayesian analysis, followed by discussions from many robust Bayesians. Several of these discussions were themselves extensive reviews of particular aspects of Bayesian robustness, so that the volume serves as a highly effective summary of the state of Bayesian robustness in 1994. Earlier reviews of Bayesian robustness were given in Berger (1984, 1985, 1990); these included considerable philosophical discussion of the approach, together with discussion of its history, including its origins in papers by Good in the '50s (Good, 1952, 1959 and 1961) and, later, by Kadane and Chuang (1978). Another good source for philosophical and historical information on Bayesian robustness is Walley (1991), although the book takes a somewhat different slant on the problem.

In the early '90s, there was an explosion of publications focusing on studying sensitivity to the prior distribution, in part because non-Bayesians often view this sensitivity to be the major drawback of Bayesian analysis. This work focused on replacing a single prior distribution by a class of priors, and developing methods of computing the range of the ensuing answers as the prior varied over the class. This approach, called "global robustness", was soon supplemented by "local robustness" techniques, which focused on studying local sensitivity (in the sense of derivatives) to prior distributions. Interest naturally expanded into study of robustness with respect to the likelihood and loss function as well, with the aim of having a general approach to sensitivity towards all the ingredients of the Bayesian paradigm (model/prior/loss). Practical implementations of robust Bayesian ideas began to appear by mid 90's (see, e.g., the previously mentioned IMS Lecture Notes volume).

In the last half of the 90's, robust Bayesian analysis shifted from being a "hot topic" to being a mature field within Bayesian analysis, with continued gradual development, but less of a sense of urgency. This change came about for several reasons. First, the initial flurry of fundamental theoretical advances naturally slowed, as will happen with any new field. Second, few Bayesians any longer questioned the need to view robustness or sensitivity as a serious issue, so that the philosophical excitement with the idea waned. Indeed, the consensus was that the time was now ripe to develop user-friendly implementations of the existing robust Bayesian methodology. The timing was not great, however, in that, coincidentally, this period also marked the explosion of interest in MCMC computational methods for Bayesian statistics. This had two serious effects on the field of Bayesian robustness. First, many of the researchers in robust Bayesian methods shifted their research into the MCMC arena. Second, the MCMC methodology was not directly compatible with many of the robust Bayesian techniques that had been developed, so that it was

[1] Based upon excerpts from the first chapter of *Robust Bayesian Analysis* (D. Rios Insua and F. Ruggeri eds.), reprinted by permission of Springer Verlag.

unclear how formal robust Bayesian analysis could be incorporated into the future "Bayesian via MCMC" world.

Paradoxically, MCMC has dramatically increased the need for consideration of Bayesian robustness, in that the modeling that is now routinely utilized in Bayesian analysis is of such complexity that inputs (such as priors) can be elicited only in a very casual fashion. It is now time to focus again on Bayesian robustness, and to attempt to bring its ideas into the Bayesian mainstream and other research fields.

About the latter goal, it is worth mentioning that Bayesian robustness is playing a relevant role in SAMO (Sensitivity Analysis of Model Output), a group interested in investigating the relative importance of model input parameters on model predictions (see Saltelli *et al.*, 2000) in many applied areas, from chemical engineering to econometrics.

The most important challenge for the field of Bayesian robustness is to increase its impact on statistical practice; indeed, to make it a routine component of applied work.

Perhaps the most important way to bring robust Bayesian methods to practitioners is to have these methods available in standard statistical software. There are several ongoing efforts to develop general Bayesian statistical software, and incorporation of robust Bayesian methodology into such software is clearly desirable.

➤ **References**

BERGER, J.O. (1984). The robust Bayesian viewpoint (with discussion). In *Robustness of Bayesian Analysis*, (J. Kadane, ed.), Amsterdam: North-Holland.

BERGER, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis.* (2nd Edition). New York: Springer-Verlag.

BERGER, J.O. (1990). Robust Bayesian analysis: sensitivity to the prior. *Journal of Statistical Planning and Inference*, **25**, 303-328.

BERGER, J.O. (1994). An overview of robust Bayesian analysis (with discussion). *Test*, **3**, 5-59.

BERGER, J. O., BETRÒ, B., MORENO, E., PERICCHI, L.R., RUGGERI, F., SALINETTI, G. and WASSERMAN L. (1996). *Bayesian Robustness*, IMS Lecture Notes - Monograph Series, **29**. Hayward: IMS.

GOOD, I.J. (1952). Rational decisions. *Journal of the Royal Statistical Society, B*, **14,** 107-114.

GOOD, I.J. (1959). Could a machine make probability judgments? *Computers and Automation*, **8**, 14-16 and 24-26.

GOOD, I.J. (1961). Discussion of C.A.B. Smith: Consistency in statistical inference and decision. *Journal of the Royal Statistical Society, B*, **23**, 28-29.

KADANE, J. B. and CHUANG, D. T. (1978). Stable decision problems. *Annals of Statistics*, **6**, 1095-1110.

SALTELLI, A., CHAN, K. and SCOTT, M. (2000). *Mathematical and Statistical Methods for Sensitivity Analysis.* New York: Wiley.

WALLEY, P. (1991). *Statistical Reasoning With Imprecise Probabilities.* London: Chapman and Hall.

BAYES AND PRICE

by Anthony O'Hagan
a.ohagan@sheffield.ac.uk

Some colleagues may be interested in historical work on Bayes. I have come across various things through correspondence with Martyn Hooper of Bridgend College in Wales, which may be well known to serious historians but were new to me.

Martyn might be called a Priceian, because he is an admirer of Richard Price, the man who presented Bayes's famous paper to the Royal Society after Bayes's death. It certainly seems that Price was a great man, and a much more significant figure in his day, but my attention was aroused by Martyn Hooper's belief that it was Price who did all the key work on Bayes's paper. If he was right, we would all have to start calling ourselves Priceans instead of Bayesians.

Martyn's studies of the life of Price meant that he was familiar with a book called *The Correspondence of Richard Price*, by Peach and Thomas, published by the University of Wales Press in 1983 (ISBN 0-7083-0819-8). (For colleagues in other countries, I should say that Price is very much a Welsh name. It is actually a corruption of Ap Rys, which in the Welsh language means son of Rys. I guess that the other common names Rice, Reese, Rees and Preece all come from the same source.)

This book includes a copy of Bayes's 1763 paper, with various extra pieces of information in annotations. If any serious historians are reading this, they should be aware of this book. The text of the paper is based on George Barnard's version published in *Biometrika* in 1958. In a footnote, the authors give an interesting short biographical note on Bayes, followed by the following:

"Bayes did not publish his work in mathematics. After his death in 1761 his relatives asked Price to go through his papers. Price did so and came upon the essay presented here which contains what he perceived to be an important contribution to the theory of probability. Price edited it, writing the introduction himself, and submitted it to the Royal Society".

Although I have heard Price described as a friend of Bayes, Martyn Hooper says that there is no evidence that the two had even met. It is not known why Price was invited to look through Bayes's mathematical notebooks.

So Price "edited" Bayes's paper. In his own introduction to the paper, Price gives some indication of the extent of the editing. He writes thus:

"The last two rules in this essay are given without the deductions of them. I have chosen to do this because these deductions, taking up a good deal of room, would swell the essay too much; and also because these rules, though of considerable use, do not answer the purpose for which they are given as perfectly as could be wished. They are however ready to be produced, if a communication of them should be thought proper. [Price did in fact publish some of this in a later paper to the Royal Society.] I have in some places writ short notes, and to the whole I have added an application of the rules in the essay to some particular cases, in order to convey a clearer idea of the nature of the problem, and to shew how far the solution of it has been carried.

...Some of the calculations, particularly in the Appendix, no one can make without a good deal of labour. I have taken so much care about them, that I believe there can be no material error in any of them; but should there be any such errors, I am the only person who ought to be considered as answerable for them."

It seems that Price added substantially to the paper. The Appendix, containing the applications to some particular cases and running to about 9 pages, was probably entirely the work of Price. He also checked very carefully Bayes's algebra and corrected at least one error. This is in "Rule 2". Actually, at the end of "Rule 1", Price writes "Thus far Mr Bayes's essay", but since it is clear that he has corrected Bayes's work in Rule 2 it is not obvious to me what this means. Perhaps Bayes's write-up of his work in the form of a proper essay ended there, and the other material was assembled and corrected by Price from other notes that Bayes had left. But that is speculation.

Martyn Hooper saw all this as evidence that Price is the one who should gain the credit for the paper. Also, he reports from another source that Price's hair

was said to have turned white overnight as he worked on Bayes's essay. It may be that Price should be given more credit than he has perhaps been up to now. He clearly did much more than just turn Bayes's paper over to the Royal Society. He saw it as a highly significant piece of work, and put a substantial amount of his own effort into it. But what is the significance of the parts that Price contributed?
Rule 1 is what we would now call Bayes' Theorem, applied to the case of a binomial sample with uniform prior. There is no doubt in my mind that this is the work of Bayes, and that the key step of using Bayes' Theorem to solve the problem of inverse probability is due to Bayes. What Price contributed was concerned with the subsequent development of Rules 2 and 3. These present alternative ways of computing the result that Bayes had given in Rule 1. Rule 1 gives the posterior probability of theta lying between any two values as the difference of two infinite series. Rule 2 gives a better series and Rule 3 gives bounds. Obviously in those days they did not have computers to calculate the incomplete beta function integral for them. The significance of Rules 2 and 3, which Price clearly saw, was that they made the practical calculation of these probabilities more feasible. Price seems to me to have been a very practical man. He went on to play a major role in the development of insurance, and for that he would have needed to calculate such probabilities, whereas Bayes may have been more content with having solved the mathematical problem and not given so much care to the computational side. Again, just speculation, but the importance of computing the incomplete beta function persisted, so that at the start of the 20th century Karl Pearson employed an army of people to produce his tables of the incomplete beta function, which he published as a massive volume.
My reading then, is that Price contributed substantially to the success of Bayes's paper, making sure that the results on computation were correct and adding an explanatory Appendix. We know that without his insight in seeing the value of Bayes's essay it would never have come to light. It is possible, too, that without Price's efforts to make it more usable and understandable the paper would not have made such an impact at the time. So I agree with Martyn Hooper that Richard Price deserves more credit, but the chief honour for having cracked the problem of inverse probability and having founded (albeit in the most tenuous way) Bayesian statistics must in my opinion stay with Thomas Bayes. I will continue to call myself a Bayesian!
P.S. Incidentally, both Bayes and Price are buried in Bunhill Fields in London.

SAILING THE BAYESIAN BOAT IN A HOSTILE SEA

by Hedibert Freitas Lopes
hedibert@im.ufrj.br

In this same section of the previous issue (March, 2001), professor Bill Bolstad guided us through his *Introduction of Bayesian Statistics* course. Those of us who read the article and constantly have nightmares when teaching Bayesian statistics, are certainly looking forward to reading his textbook (still under preparation). He also presents a table that shows that during the seventies and eighties no more than 6% of JASA articles were clearly Bayesian by their title, with a considerable increase to about 15% in the late nineties. Most of us would agree that Markov Chain Monte Carlo (MCMC) methods played, not surprisingly from a practical viewpoint, a very important role in facilitating the lives of Bayesian researchers and practitioners all over the world, either when trying to solve a problem from a purely Bayesian perspective or when trying to combine Bayesian and frequentist techniques.

In research and application levels, the use of Bayesian methods has been steadily increasing and has found greater acceptance. Successful examples of this wide reception include the existence of this society, its world meetings, the Valencia meetings, several regional chapters (Brazil, Chile, etc.) and other meetings, such as the six biannual Brazilian meetings of Bayesian statistics and the upcoming first Latin

American Meeting of Bayesian Statistics to be held in São Paulo in February 2001 (`www.est.ufmg.br/cobal`, with the English version under preparation). All these meetings are honored with several interesting and important applications, from archaeology to financial time series to genetic expression analysis. However, when it comes to teaching, we are still behind several years if not decades. Statistical departments around the world are still heavily frequentist and, apart from demographic reasons, it will still take many years before Bayesian statistics will escape chapters 15, 16 (or even 24) of most textbooks on introductory statistics where it is seen as and taught to be simply another statistical tool or technique. This is why it is an appropriate time to thank professor Morris DeGroot for his brilliant book on Probability and Statistics, published more than 30 years ago, that has certainly inspired many other authors, some of which are members of this society.

Unfortunately, the several difficulties of teaching Bayesian statistics are exacerbated in developing countries, such as Brazil, by the lack of inexpensive or free computing resources. Our best universities are public, where the students do not pay a penny for their education and most of them would withdraw from school should they have to pay any tuition or fee. Therefore, ingenious ways of overcoming such difficulties are always on the top of the list. Just to give a brief and short personal example, if I may, when returning from my PhD studies last June I saw myself with the duty of teaching Statistical Computing to master students in the department of statistical methods at my university (Federal University of Rio de Janeiro) and later on to undergraduate students, also in the statistics department. Coming from Duke, I was already used to several toys, such as Splus, Matlab and Minitab, all of which are really expensive for our students to purchase. The point of the soap opera is to show that we commonly overlook these problems when we focus mostly on research.

The search for free Bayesian teaching ammunition and weapons already has some interesting starting points. For instance, the R project, Splus youngest cousin (see Ribeiro and Brown's article on the previous issue of this bulletin), is free and open source, and more importantly it runs most of the Splus codes without major changes (`www.r-project.org`). It certainly helps solve a great number of statistical problems faced by undergraduate and graduate statistics students. Linear models, multivariate and time series analysis and econometrics, to name a few, can almost entirely be taught from a Bayesian perspective by using statistical packages such as R and BUGS (Bayesian inference Using Gibbs Sampling - information available at `www.mrc-bsu.cam.ac.uk/bugs`). Several of our master students in recent years have extensively based their research projects on WinBugs coding. Also, Tony O'Hagan's First Bayes software is of great assistance when explaining to undergraduate students the intuitive nature of the Bayesian paradigm. It is becoming common practice, among Bayesians, to publicly share computational routines and software. I believe the next couple of decades will witness a uniformity across countries unconditional of their economical status, mainly due to the velocity with which (Bayesian) information is being spread. In order to teach good statistics courses we, responsible Bayesians, will have to learn how to combine important existing Bayesian and non-Bayesian techniques in order to best prepare the future generations. Once again, as professor Bolstad said well *"If we get the future scientist thinking Bayesian from the start, they will be doing Bayesian analysis when they write up their research"*. I would simply add that they, future scientists and decision makers, will be better prepared to impartially judge the validity of Bayesian proposals, should they be in a scientific paper, a new clinical trial scheme, or even an economical shock, something we Brazilians are accustomed to.

BAYESIANS IN IRELAND

by Cathal Walsh and Simon Wilson
walshc@tcd.ie
simon.wilson@tcd.ie

In the last decade the Bayesian brush has been steadily painting the Emerald Isle its own colour. Some of those doing the painting have philosophical grounds for choosing their methods, whereas the majority are pragmatists, and subscribe because the methods work for them in practice. A search for current Bayesian activity was undertaken in Ireland, and reasonable success was encountered.

A famous Irish song contains the words "Just to see again the moon rise over Claddagh and to watch the sun go down on Galway Bay." Unfortunately it's been impossible to track down any Galway Bayesians on whom the sun may some day set, but in the University of Limerick, just a short distance south of Galway, Don Barry flies the flag.

Don is a professor at the University of Limerick and works on Bayesian models for smoothing. A recent work of his was entitled "Bayesian analysis for less smooth departures from polynomial regression models," published in *Communications in Statistics* last year. He has a number of PhD students working in the area of Bayesian models for partitions: Victoria Livingstone has developed models for partitioning of studies in a meta-analysis; Avril Hogan is applying partition models to the analysis of spatio-temporal variation in disease rates; Claire Jordan is using partition models in the context of classification and regression trees.

Other Bayesian practitioners have found their home amongst the ancient cobblestones of Trinity College Dublin. A critical Bayesian mass has not yet been formed - the various parties have been scattered through different departments. Some have even convinced the established practitioners of the usefulness of their methods!

In the Department of Statistics, Brendan Murphy, Cathal Walsh, Matt Whiley and Simon Wilson are challenging the balance of 'power'. The Chair is happy to use the Bayesian methodology, provided it is appropriate for the application.

The Professor of Statistics at Trinity, John Haslett is working with Matt and Simon on a model based statistical approach to spatial palaeo-environmental reconstruction. While John wouldn't consider himself a "Bayesian" he is happy to use the methodology for a particular project. This application is a prime example of a situation in which the natural structure of the problem is best described through a Bayesian hierarchical framework.

In describing the problem, John outlined that in the past, quantitative procedures, often referred to as transfer functions methods, have been available to make palaeo-environmental reconstructions from different types of fossil assemblages including those of pollen, diatoms and chironomids. Such models are calibrated on modern data, involving a wide variety of environments at some thousands of sites across Europe. Current methods do not address many of the important spatial features of such calibration data sets, and the models are thus inadequate for exploring detailed research hypotheses. The objective in this situation is to use modern computationally intensive statistical methods based on Bayesian hierarchical models.

Brendan Murphy returned to take up a post in Ireland in recent years, having completed a PhD at Yale. His current research interests are in Bayesian robustness and latent variable problems (in particular mixture models). A publication with Hartigan on "Inferred Probabilities" is soon to appear.

Cathal Walsh is one of Simon's former graduate students, and was appointed last year as a lecturer at Trinity College, having spent some time as an investment analyst. His postgraduate research was on Bayesian modelling of short fatigue crack growth and coalescence, a reliability problem in the applied materials engineering field. The particular application is detailed in the volume *Frontiers in Reliability* (Basu, Basu and Mukhopadhyay eds.), World Scientific. Cathal has continuing interest in this area, but has also done some work on problems with financial and medical applications, which follow from some work started during his time in the commercial world.

The analysis of the spatial interaction of cracks in metal is a project which a PhD student, Eleisa Heron, is currently working on.

Matt Whiley is a postdoctoral researcher and is currently looking at parallel algorithms for Bayesian inference of spatial gaussian models with Simon and John. The current methodology is highly computationally expensive, which places an effective limit upon the number of spatial locations we can consider. The aim of this work is to utilise the extra computational power available on parallel computers and thus allow the extension of this methodology to larger problems. He is also working with John on the reconstruction of past climates.

Simon Wilson was one of the first Bayesian influences on the Department of Statistics at Trinity. He has engaged a number of graduate students in different projects over the years, ranging from estimating the number of bugs in computer software, through image segmentation to cracks in metal. At present, Simon is a key figure in image restoration work, which he is pursuing with colleagues across campus in Electronic Engineering. He has two PhD students, Breedette Hayes and Giorgios Stefanou working in spatial and temporal modelling and image recognition and retrieval respectively.

Working in the Dept. of Electronic and Electrical Engineering, Anil Kokaram heads up the "Signal Processing and Media Applications Group." On his election to Fellowship of the University this year, he was acknowledged as one of the world's leading experts on the digital restoration of motion pictures. His 1998 book on that theme has become the seminal work. His work in the broader field of multimedia signal processing is widely respected. He is co-ordinator of an EU-funded Research Training Network in Information Retrieval (MOUMIR) and is also involved with the EU Project BRAVA which deals with Restoration of Archived Video. For more information see `www.mee.tcd.ie/~ack` and `www.mee.tcd.ie/~sigmedia`.

Also of the same Department, Anthony Quinn has used Bayesian techniques extensively for the identification of parametric signal and image models. In particular, he is interested in the way that marginal inference strategies engender Ockham's Razor. These ideas have been used to develop regularized model identification strategies for spaces of models of varying complexity. He is involved in a European research initiative to develop fast Bayesian approaches to mixture model identification for use in decision feedback. This work is being undertaken with two postdoctoral researchers, Ludvik Tesar and Ladislav Jirsa, who have joined Anthony from a partner laboratory in Prague. In the image analysis domain, Anthony has worked with his student, Ed Clark, on the development of design principles for Markov Chain Monte Carlo samplers in image segmentation. Anthony has provided a Bayesian contribution in *Signal Analysis and Prediction* (Prochazka et al. eds.), Birkhauser, 1998.

Apart from the on campus statisticians and engineers, Bayesian thinking has influenced others in Trinity College Dublin. For several years Alan Kelly, a lecturer in Biostatistics in the Department of Community Health and General Practice has been employing small-area Bayesian spatial models to analyse Ward level mortality data in collaboration with several regional Health Boards - mainly in Ireland but also in Wales. Currently, Alan is working with the cancer registries in Northern Ireland and the Republic of Ireland to complete an all Ireland small-area analysis of major cancers for adults and Leukaemia for children. Also now working on developing spatio-temporal models for 20 years of Irish mortality data attempting a fresh look at excess mortality during 'winter' months. At the European level, Alan is a partner in a 8 country EU-funded project (EUROHEIS) to analyse small-area mortality/morbidity patterns in conjunction with information on area deprivation. A recent contribution by Alan was "Case studies in Bayesian disease mapping for health and health service research in Ireland" in *Disease mapping and risk assessment for Public Health* (Lawson et al. eds.), Wiley, 1999.

All in all, there is plenty of activity under the heading "Bayesian" in Ireland. Rest assured there are many others who are almost convinced. At present it is polite to talk about such people as "Likelihood with a uniform prior." It has been useful to examine the areas in which Irish statisticians work, and encouraging to see the growth of of Bayesian methods in this state.

BAYESX: MCMC SOFTWARE FOR SEMIPARAMETRIC REGRESSION AND GAUSSIAN DAGS

by Nicky Best
n.best@ic.ac.uk

➤ Introduction

BayesX is a software tool for Bayesian semiparametric regression written by Stefan Lang and Andreas Brezger at the University of Munich. The program allows estimation of a wide range of models within the class of generalized linear and generalized additive mixed models (GLMMs and GAMMs; see e.g. Clayton, 1996; Fahrmeir and Lang, 2001), the latter via a variety of flexible smoothing priors for nonlinear temporal and other quantitative covariate effects. BayesX is also able to estimate varying coefficient models (Hastie and Tibshirani, 1993). That is, interactions in which the effect of a particular covariate is allowed to vary in a smooth, nonlinear fashion over the domain of a second (metric) covariate. Spatial covariates may also be estimated using Markov random field priors. The other main feature of BayesX is the ability to estimate the structure and model parameters of Gaussian DAGs (Directed Acyclic Graphical models; Geiger and Heckerman, 1994). The software uses MCMC inference methods for estimation.

So, why am I writing a review of BayesX? I first came across the software when attending some lectures and practical sessions given by Leo Knorr-Held as part of a short course on longitudinal modelling at Imperial College earlier this year. Leo persuaded me to write a review of the software for this Bulletin (probably to encourage me to use some software other than WinBUGS!) and so this seemed like a good incentive for me to find out more about the program's capabilities. My comments are therefore very much first impressions written from the perspective of a novice user of the BayesX software. My experience is limited to having experimented with some simple example datasets, and I have yet to use the program "in anger". I therefore apologise to the authors if my comments fail to do justice to all the features in software.

➤ Basics: platforms, where to get it, what you get, etc.

BayesX is only available for Windows. The most recent version (0.6) can be downloaded for free from www.stat.uni-muenchen.de/~lang/bayesx/bayesx.html. It comes as a zipped executable and is straightforward to install. The installation file includes a separate manual (see below), some Splus functions for plotting the output from BayesX, and two example datasets. The Windows interface is relatively basic, essentially providing the user with a command window in which to type BayesX statements for loading and manipulating data, specifying regression models and driving the MCMC simulations. A log file (review window) records all the commands entered and allows for easy recall and editing of previous statements, and output is written to a third window that can be saved as a text or RTF file. A useful feature is that the program may also be run in batch mode.

I haven't seen anything written in the documentation about the processor or memory demands of the software, but for the simple (and rather small) datasets I have analysed, my 333MHz Pentium laptop with 96MB RAM is more than adequate.

➤ Documentation and support

BayesX comes with a hefty manual of some 111 pages, available as a separate postscript file. The authors have clearly put a lot of effort into preparing the documentation; however there is no real shortcut to getting started with BayesX other than working your way systematically through the manual. The two example datasets are used to illustrate the BayesX command syntax and various features available in the software. In a future version it would be nice to see some further examples, perhaps organised in a separate section to make for easier reference. The manual does contain a concise but extremely useful review of the methodology underlying Bayesian GAMMs, and of the MCMC estimation techniques implemented in BayesX, plus pointers to the relevant literature.

There is currently no on-line help, although a 'help' button appears on the menu bar, indicating that the authors may

be planning to include this feature in a future version. This would certainly be a valuable addition to the software, or at least the provision of a 'crib sheet' in the manual covering syntax and usage of the most common `BayesX` commands.

➤ **Key features**

GAMMs

`BayesX` supports Gaussian, Poisson, binomial, multinomial and ordered categorical response data, and can be used to fit regression models with canonical (or, in some cases, probit) link function and a parametric, semiparametric or nonparametric predictor. The key feature of the software is the range of prior distributions available for estimating nonlinear covariate effects. These include first and second order random walks, P-splines (of fixed degree) with first or second order random walk penalty, smoothing splines, time-varying seasonal effects, Markov random fields (MRF) for spatial (geographical) effects and exchangeable Gaussian random effects. Any of these priors may also be used to specify the distribution of the effect modifier in a varying coefficient (interaction) term. `BayesX` also provides facilities for reading geographical map files (either in the form of an adjacency matrix or as a list of polygon coordinates that the program then uses to create the adjacency matrix) needed by the MRF prior for spatial effects. To give a flavour of the straightforward command syntax needed to specify a relatively complex GAMM in `BayesX`, an example is shown below of part of the `BayesX` statement for fitting a semiparametric regression model for a Gaussian response `Y` regressed on covariates `X` (as a fixed effect), `AGE` (as a second order random walk) and geographical location `AREA` (as a spatial MRF effect plus an exchangeable random effect):

```
Y=X+AREA(spatial,map)
+AREA(random)+AGE(rw2)
```

MCMC algorithms

A second major strength of the `BayesX` software is the computational efficiency of the MCMC algorithms used. The authors exploit the fact that for the specific classes of models implemented in `BayesX` there exist efficient block updating schemes that can considerably improve mixing and convergence. For Gaussian responses, the joint full conditionals are multivariate normal with band precision matrices that can be sampled efficiently using Cholesky decomposition for band matrices (Rue, 2001). An algorithm for reordering the regions in a geographical map to create a banded precision matrix for the spatial random effects is also implemented. Binomial, multinomial and ordered categorical responses with probit link also lead to multivariate normal full conditionals. For non-Gaussian, non-probit-link regression models, `BayesX` uses a modified version of Gamerman's (1997) weighted least squares proposal for fixed effects and a Metropolis-Hastings algorithm with conditional prior proposals developed by Knorr-Held (1999) for random effects. Block updating is possible with these algorithms although the blocksize depends on the dataset and must be specified by the user (for models with multivariate Gaussian full conditionals, `BayesX` automatically updates all parameters as a single block). There appear to be a few teething problems with the user-specified block size (at least, I managed to cause the program to crash on a number of occasions when specifying what I thought was a sensible block size!).

Gaussian DAGs

As mentioned in the introduction, `BayesX` can also be used to estimate the dependency structure and model parameters for a set of variables represented by a Gaussian DAG. The user must specify the variables of interest, and `BayesX` then uses a reversible jump (RJ) MCMC algorithm to add, switch or delete directed edges between each pair of variables and estimate the corresponding set of model parameters assuming a normal linear regression model for each variable given its 'parents' in the current DAG. The appropriate sampling distributions are determined using the same factorization and Markov properties of DAGs that underlie the way in which the `WinBUGS` software calculates full conditional distributions. However, whereas `WinBUGS` is only concerned with sampling from a DAG with a fixed structure pre-specified by the

user, the dag option in BayesX allows *learning* about the model structure (i.e. presence or absence of directed links) via the RJMCMC estimation algorithm.

➤ **Comparison with WinBUGS**

Although this is a review of BayesX, I decided to take the opportunity for a quick commercial break to make a few comments on the WinBUGS software (Spiegelhalter et al, 2000) and discuss how this compares with BayesX for estimating the classes of models implemented in the latter. I imagine that most readers have at least heard of WinBUGS, and a number will have used it. For anyone not familiar with WinBUGS, it is currently the most general-purpose software tool available for Bayesian inference using MCMC sampling methods. This contrasts with most other Bayesian software (including BayesX) which tends to be application-specific or at least restricted to certain fairly specific classes of model. There is an obvious trade-off between flexibility and efficiency that is well illustrated by comparing WinBUGS and BayesX. My first attempt at using BayesX for my own data failed when I attempted to fit a Poisson regression model with spatial random effects to counts of disease cases and population denominators in a set of geographically contiguous areas: I wanted to include the population in each area as an offset but the BayesX syntax does not currently allow this. No doubt this is a simple enough modification for future versions of BayesX, but demonstrates difficulty that software authors face in trying to anticipate the full range of models that users may want to fit.

The offset problem for my disease mapping data could be side-stepped by assuming a binomial rather than Poisson likelihood. In this case, fitting the same model in WinBUGS and BayesX yielded very similar results in terms of program speed, convergence, mixing, autocorrelation functions and posterior summaries. Estimation of a binomial regression model with a second order random walk smoothing prior for the effect of age also yielded very similar results from both BayesX and WinBUGS, although the WinBUGS posterior samples for the age effects were more highly autocorrelated than those generated by BayesX. In principle (although I haven't tried) it should also be possible to fit the other smoothing priors available in BayesX using WinBUGS. However model specification is likely to require some clever use of the WinBUGS language and other 'tricks', since, with the exception of the MRF spatial prior (implemented in WinBUGS using the car.normal distribution) and the second-order random walk (illustrated using the ICE example in volume II of the on-line WinBUGS examples manual) none of these distributions are available directly in WinBUGS. In the context of spatial MRF priors, Knorr-Held and Rue (2001) also demonstrate that in situations with sparse data, single-site updating schemes (such as those currently implemented in WinBUGS) may struggle to fully explore the tails of the posterior distribution for the random effects. The block updating algorithms proposed by Knorr-Held and Rue show better mixing in such situations, although note that these are not the same as the block updating schemes currently implemented in BayesX.

RJMCMC algorithms are not available in WinBUGS, and so it is not possible to implement model selection for Gaussian DAGS in WinBUGS.

➤ **Should you use BayesX?**

The bottom line is "yes" if you want to fit any of the GLMMs, GAMMs or Gaussian DAGs currently supported by BayesX. The software is reasonably easy to use, implements some very powerful and efficient MCMC algorithms and provides some helpful graphics functions for visualisation of output in Splus. I would certainly recommend BayesX to colleagues and students in my department and I congratulate Stefan Lang and Andreas Brezger for making this software available. However if your application involves additional complexity such as covariate measurement error, informative missing data, prediction, non-standard prior distributions etc. then the flexibility of the more general-purpose WinBUGS software or the specificity achieved by writing your own code remain your only options. The ingenuity of the growing number of users of Bayesian MCMC methods to devise new

and increasingly more complex models for their applications presents a continuing challenge to all Bayesian software developers to provide efficient, robust and user-friendly software tools to encourage and support the widespread use of Bayesian methodology.

➤ **References**

CLAYTON, D.G. (1996) Generalized linear mixed models. In W.R. Gilks, S. Richardson and D.J. Spiegelhalter (eds.), *Markov chain Monte Carlo in practice*. London, Chapman and Hall, 275–301.

GAMERMAN, D. (1997) Efficient sampling from the posterior distribution in generalized linear models. *Statistics and Computing*, **7**, 57–68.

GEIGER, D. and HECKERMAN, D (1994) Learning Gaussian networks. *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, San Francisco, Morgan Kaufmann, 235–243.

FAHRMEIR, L. and LANG, S. (2001) Bayesian inference for generalized additive mixed models based on Markov random field priors. *Applied Statistics (JRSS C)*, **50**, 201-220.

HASTIE, T. and TIBSHIRANI, R. (1993) Varying-coefficient models. *Journal of the Royal Statistical Society, B*, **55**, 757–796.

KNORR-HELD, L. (1999) Conditional prior proposals in dynamic models. *Scandinavian Journal of Statistics*, **26**, 129–144.

KNORR-HELD, L. and RUE, H. (2001) *On block updating in Markov random field models for disease mapping*. Technical report, University of Munich (`www.stat.uni-muenchen.de/~leo/publikationen.html`).

RUE, H. (2001) Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society, B*, **63**, 325–338.

SPIEGELHALTER, D.J, THOMAS, A. and BEST, N.G. (2000) *WinBUGS Version 1.3 User Manual*. Mrc Biostatistics Unit, Cambridge and Imperial College, London. (`www.mrc-bsu.cam.ac.uk/bugs`).

NONPARAMETRIC BAYESIAN SURVIVAL MODELS: AN ANNOTATED BIBLIOGRAPHY

by Jaeyong Lee
`leej@stat.psu.edu`

Although nonparametric Bayesian survival models appeared more than two decades ago, the area is still new and active. As the bibliography presents itself, more things need to be done than what has been done. I hope that this bibliography gives a good summary of history and provides a good starting point to students and researchers who want to contribute to this exciting area.

The bibliography is categorized into general theory, asymptotic theory, posterior computation and other issues with one more section devoted to general nonparametric priors.

➤ ***General Nonparametric Priors***

The following two papers are not directly related to survival analysis, but they are cornerstones of the nonparametric Bayesian literature and the area of nonparametric survival analysis is not an exception.

• T.S. FERGUSON (1973). **A Bayesian Analysis of Some Nonparametric Problems.** *The Annals of Statistics*, **1**, 209-230.

This is the seminal paper for the area of Bayesian nonparametric statistics. It broke the Bayesians long silence on nonparametric statistics. Dirichlet processes are defined and their basic properties, conjugacy and discreteness, are proved. It is a must read.

• K.A. DOKSUM (1974). **Tailfree and Neutral Random Probabilities and Their Posterior Distributions.** *Annals of Probability*, **2**, 183-201.

Processes neutral to the right and tailfree processes are introduced in a general space. A simple characterization of processes neutral to the right on the real line is introduced, i.e., if F is a process neutral to the right, the following characterization holds:

$$F(x) = 1 - e^{-Y(x)}, \qquad (1)$$

where Y is an independent increment process. This characterization is later used in survival models with Y interpreted as a cumulative hazard function.

➤ ***General Theory***

• V.J. SUSARLA AND J. VAN RYZIN (1976). **Nonparametric Bayesian Estimation of Survival Curves from Incomplete Observations.** *Journal of the American Statistical Association,* **71**, 897-902.

Dirichlet process is used as the prior for the distribution function. Only the Bayes estimator of the survival curve is derived. Since the class of Dirichlet processes is not conjugate with censored data, it was a natural consequence. It has also been noted that in the limit the Bayes estimator converges to the Kaplan-Meier estimator.

• J.D. KALBFLEISCH (1978). **Non-parametric Bayesian Analysis of Survival Time Tata.** *Journal of Royal Statistical Society B* **40**, 214-221.

The author describes a Bayesian treatment of the celebrated Cox's proportional hazard model with a gamma process prior on the baseline cumulative hazard function.

• T.S. FERGUSON AND E.G. PHADIA (1979). **Bayesian Nonparametric Estimation Based on Censored Data.** *The Annals of Statistics* **7**, 163-186.

The authors worked with a large class of processes neutral to the right as the prior class for the distribution and used the characterization of process neutral to the right (**??**). Under right censoring, the class of neutral to the right process is a conjugate prior class with right censored data and the posterior distribution is obtained.

• R.L. DYKSTRA AND P. LAUD (1981). **A Bayesian Nonparametric Approach to Reliability.** *The Annals of Statistics*, **9**, 356-367.

The authors define an extended gamma process and use it as a prior process for the hazard rate function. It is a main difference from other research works that the authors work with the hazard rate function rather than the cumulative hazard function. A natural consequence is that it allows only increasing hazard rate functions.

• L.M. BERLINER AND B.M. HILL (1988). **Bayesian Nonparametric Survival Analysis.** *Journal of the American Statistical Association*, **83**, 772-779.

The authors of this paper discuss a notably different approach than the other papers introduced here. They emphasize the role of prediction and derive prediction probability based on exchangeability.

• N.L. HJORT (1990). **Nonparametric Bayes Estimators Based on Beta Processes in Models for Life History Data.** *The Annals of Statistics* **18**, 1259-1294.

It is this paper that the prior is directly put on the cumulative hazard function rather than using the characterization (**??**) for the first time. The author first describes the beta process in discrete time model and shows how the beta process arises naturally. Then, he extends it to continuous time model. To obtain the posterior distribution of the beta process, the author works with a general class of Lévy process priors on the cumulative hazard functions and obtained the posterior distribution with Lévy process priors and subsequently beta processes. This is a must read.

• A.Y. LO (1993). **A Bayesian Bootstrap for Censored Data.** *The Annals of Statistics*, **21**, 100-123.

This paper has two important contributions. One is the generalization of Rubin's Bayesian bootstrap to censored data setting. The other is independent construction of beta process, which the author termed beta-neutral process. The derivation of the beta-neutral process is simple and quite different from that of Hjort (1990) and the derivation uses two gamma processes for censored and uncensored observations.

• S. WALKER AND P. MULIERE (1997). **Beta-Stacy Processes and a Generalization of the Pólya-urn Scheme.** *The Annals of Statistics*, **25**, 1762–1780.

The authors introduce a class of prior processes on the space of distribution functions and call it beta-Stacy processes. The cumulative hazard function of this process is the beta process.

• Y. KIM (1999). **Nonparametric Bayesian Estimators for Counting Processes.** *The Annals of Statistics*, **27**, 562-588.

The author adopts the approach taken by Hjort (1990) and substantially extends it to the much more general framework of multiplicative

intensity model which has shown its success for decades. Previously, all the development has been based on the moment generating function representation of Lévy process. It is in this paper that the representation of Lévy process as an integral of a Poisson measure can be fruitfully used in the Bayesian nonparametric survival models.

➤ ***Asymptotic Theory***

• J.K. GHOSH AND R.V. RAMAMOORTHI (1996). **Consistency of Bayesian Inference for Survival Analysis with or without Censoring.** in *Analysis of Censored Data* (eds: Koul, H.L. and Deshpande, J.V.)

The posterior consistency was one of central theoretical themes in nonparametric Bayesian research for there are many examples that innocent-looking priors have inconsistent posteriors. This paper deals with the posterior consistency issue with censored data and Dirichlet process prior on the distribution function.

• Y. KIM AND J. LEE (2001). **On Posterior Consistency of Survival Models.** *The Annals of Statistics*, To appear.

The authors report an interesting result that not all the processes neutral to the right have consistent posteriors and characterize a set of sufficient conditions for posterior consistency. It is also shown that all the popular prior processes, Dirichlet processes, gamma processes, beta processes, have consistent posteriors.

➤ ***Posterior Computation***

• H. DOSS (1994). **Bayesian Nonparametric Estimation for Incomplete Data Via Successive Substitution Sampling.** *The Annals of Statistics*, **22**, 763-1786.

This is the first attempt to use Markov chain Monte Carlo (MCMC) for posterior computation in survival models, although the computation scheme proposed in this paper is not popular in Dirichlet process computation. However, it shows versatility of Bayesian computation with general censoring scheme, for complication of left censoring does not pose any more challenge than right censoring only. To generate an MCMC sample for the posterior distribution, Sethuraman's representation for a Dirichlet process is used.

• P. DAMIEN, P.W. LAUD, A.F.M. SMITH (1996). **Implementation of Bayesian Nonparametric Inference Based on Beta Processes.** *Scandinavian Journal of Statistics*, **23**, 27-36.

A Markov chain Monte Carlo implementation for beta processes based on a Poisson approximation of Lévy process is discussed.

• P.W. LAUD, P. DAMIEN, AND A.F.M. SMITH (1998). **Bayesian nonparametric and covariate analysis of failure time data.** in *Practical Nonparametric and Semiparametric Bayesian Statistics* (eds: Dey, D., Müller, P. and Sinha, D.)

A Markov chain Monte Carlo implementation for Cox's proportional hazard model with beta process prior is described.

➤ ***Other Issues***

• D.G. CLAYTON (1991). **A Monte Carlo Method for Bayesian Inference in Frailty Models.** *Biometrics*, **47**, 467-485.

This is the first example in the area of nonparametric survival models that the Bayesian approach is practically more usable than the frequentist counterpart. One of the inventors of the frailty models offers a Bayesian solution to a nontrivial computational problem.

• D. GAMERMAN (1991). **Dynamic Bayesian Models for Survival Data.** *Applied Statistics*, **40**, 63-79.

The author shows how the idea of dynamic linear model can be used to construct a correlated prior process.

• D. SINHA (1993). **Semiparametric Bayesian Analysis of Multiple Time Data.** *Journal of the American Statistical Association*, **88**, 979-983.

This paper gives a Bayesian analysis of proportional intensity model of multiple event data.

• D. SINHA AND D.K. DEY (1997). **Semiparametric Bayesian Analysis of Survival Data.** *Journal of the American Statistical Association*, **92**, 1195–1212.

This paper gives an overview of this active research area and discussion of fairly variety of topics.

NEWS FROM STUDENTS

by M. Eugenia Castellanos and Javier Morales
me.castella@umh.es
j.morales@umh.es

In this issue we present three papers. In the first one, a young statistician, Beatriz Peñaloza, tell us her experience working in a financial institution. We include also an employment offer for students with a background in statistics and we conclude the Student's Corner with the abstract of the thesis of Dr. Raquel Montes.

Beatriz Peñaloza Nyssen
BBVA - Bancomer
Subdirector Risk Management
Mexico City, Mexico
penaloza@data.net.mx

When I finished my graduate studies, my working perspectives were related to some statistical consulting business, perhaps doing something on marketing research (Nylsen, IBOPE, etc.). However, by chance I contacted one of the greatest banks in Mexico, Bancomer, and its Department of Consuming Risk Administration. They were interested in developing methodology to calculate the expected loss. My master degree, which treated of Bayesian Linear Classification, interested them and I started to work in the bank. By then, they only worked on screening and risk control on credit cards. Later, we incorporated other products as personal loans and mortgages, working even on profitability measures. The application of this work was immediate: to decide on credits by minimizing risk and maximizing profits on portfolios.

In Bancomer, they worked with discrimination models to accept clients and behavior models for the follow up of consumers (credit scores and behavior scores); these models had been developed by a consulting business on the United States. When I started there, my job consisted in the validation of these models as decision tools, and studying their forecasting value at the portfolio. At the same time, some models on expected loss were developed.

In Mexico there are clear patterns of economic cycles, as in consumption behavior. So, I became studying time series, with the ultimate goal of using the information from the past to be more effective on forecasting models. Very soon I perceived the limited power of classical time series. Dynamic models would be better. However, the models developed by the consulting business at the US did not take into account dynamic variables; perhaps because they relied on the economic stability of their country.

Bancomer gave me support to continue my studies on a Master on Statistics at Warwick University (Britain). I wanted to study Linear Dynamic Models in order to apply them on developing a good methodology for forecasting. Warwick University, as well as my advisor Prof.Jeff Harrison, fulfilled my expectations. There, I obtained the Master Degree on Statistics, with the work entitled "Dynamic Volatility Models".

In 1999, some models on risk and profits (at the level of client) were developed by Bancomer. These models offered the bank an integral knowledge about the client. This project was not implanted, but an integral knowledge on the client was the objective of all starting policies on and strategies for screening, always taking into account the observed risk an profits.

Now (since March 2001), I am at the Development Department of the BBVA-Bancomer, working on models for measuring the risk on loans for small and medium business. I am interested in continuing my post-graduate studies on Credit Risk Measures. My objective is to develop new measuring techniques and models, adjusted by economic factors (in countries with high volatility), as well as the measurement of the bias due to use information from different economic scenarios. My proposal is based on Bayesian techniques.

Bayesian Statistics allows to incorporate previous experience of credit risk models. Moreover, inference problems can be solved as decision ones. In practice, using Bayesian techniques is not an easy task when creating efficient models, specially because of the "fear to the unknown". However, though I have never formally applied the Bayesian methodology, loss and utility functions have been of great value to formulate my problems on decisions about the measurement method to use.

Statistics consists not only in

transforming data into information, but also in creating knowledge by developing new methods and models.

PhD Studentship in Biostatistics/Pharmacoepidemiology at the University of Southern Denmark, Odense

Project Title: Statistical Methods for the Analysis of Pharmacoepidemiological Databases.

One of the main methodological developments in pharmacoepidemiology is the emergence of large databases with more or less complete capture of individual drug use and clinical outcomes for large populations. Odense, for example, is home of the Odense Pharmacoepidemiologic Database (OPED) where all prescriptions delivered to pharmacies in the County of Fyn are registered since 1991.

Until now, there have been little efforts to develop statistical methods for the analysis of such databases. Most medical research based on pharmacoepidemiological databases makes use of traditional statistical methods borrowed from other fields, e.g. the analysis of cohort and case control studies in epidemiology. However, it is sometimes questionable, whether such methods can be used without adaptation, because pharmacoepidemiological databases show some specific features. For examples, some investigations are subject to length biased sampling due to the fact, that only events in a certain time window can be observed. Other applications require the true treatment or disease status to be reconstructed from the observed events. This is today often based on rather arbitrary decision rules, which may introduce a systematic bias.

In a previous Ph.D. project methods to estimate incidence and prevalence based on pharmacoepidemiological databases have been investigated. The current project should focus on the analysis of event patterns. It should cover both a critical review of simple ad hoc methods to evaluate basic characteristics, for example the mean duration between events, as well as the development of more sophisticated approaches to describe the distribution of the time between events by modeling the individual course as well as the heterogeneity among subjects. The latter might be extended to include latent processes, if the events observed allow only a partial reconstruction of the process of interest.

The position will be attached to the Biostatistics group (`www.isd.sdu.dk/biostat`) at the Department of Statistics and Demography (`www.sdu.sk/tvf/statdem`) and to the Research Unit of General Practice (`www.sdu.dk/health/Research/Units/Genpract.htm`). The Ph.D. student should participate in several applied projects based on OPED and similar databases and conducted at the Research Unit or at the Institute of Clinical Pharmacology. There is a possibility to work some months prior to the enrollment on a research assistance level in order to prepare the final application to the Faculty. The position should be filled during the year 2001.

Applications are invited from suitable candidates with a background in statistics. Candidates should have a degree in statistics, biostatistics, mathematics, epidemiology or similar subjects. They should be interested in working on the boundary between application and theory, should have an interest in methodology and should have good computer skills.

Further information can by obtained from Werner Vach, Professor of Medical Statistics, Department of Statistics and Demography, University of Southern Denmark, Sdr. Boulevard 23a DK-5000 Odense C (email: `werner.v@statdem.sdu.dk`) or Jakob Kragstrup, Professor of General Practice, Institute of Public Health, University of Southern Denmark, Winslowparken 19 DK-5000 Odense C (email: `jkragstrup@health.sdu.dk`)

Applications (including a cv) should be sent to Werner Vach.

Dr. Raquel Montes
ESCET. Rey Juan Carlos University
`r.montes@escet.urjc.es`

Optimal design of two-stage screens: A Bayesian approach.

Advisor: Dr. David J. Laws

The quality of an item may be described by a performance variable Y, so that an item is

accepted, as conforming, if Y belongs to a certain specification region C_Y and rejected, as non–conforming, otherwise. Assessment of the performance variable is error–free but often very expensive or difficult to determine. If available, a cheaper and easier to measure, correlated screening variable X may be used instead of (one–stage screens) or as well as (two–stage screens) the performance variable Y.

In this thesis we present a Bayesian approach to the development of simple two–stage screen designs for different model assumptions.

We develop an alternative screening plan in which the items are sequentially submitted for acceptance inspection. This sequential procedure will allow us to (1) update the posterior distribution of the unknown parameters as more data become available and (2) minimise the total expected cost of screening all the items of interest.

We also consider the case in which the relationship between X and Y is simply defined by expressing the performance variable as a function of the screening variable, that is $Y = z(X)$ for a certain unknown function $z(\cdot)$. General Gaussian Processes theory presents an attractive way of expressing prior beliefs about the function $z(\cdot)$ and we show how, in this context, a combination of analytical and numerical (MCMC) methods may be used for making inference about the posterior predictive distribution of $Y \mid X$.

Finally, the idea of accepting/rejecting items is extended to the general situation in which articles can be allocated to one of N different groups or categories. Under a particular set–up based on misclassification costs, we develop one–stage and two–stage categorical screening procedures.

NEWS FROM THE WORLD

by Antonio Lijoi
lijoi@unipv.it

✻ *denotes an ISBA activity*

➤ Events

Australasian Region of the International Biometrics Society and New Zealand Statistical Association Joint Conference 2001. *December 10-13, 2001, Christchurch, New Zealand.*

Details on the conference can be found at `nzsa.rsnz.govt.nz/Conference/home.htm`.

As part of this conference there will be a Workshop on Bayesian Methods (see details at `www.stats.waikato.ac.nz/workshop.html`) given by Dr. Bill Bolstad, Dr. James Curran, of the University of Waikato and Dr. Martin Upsdell of AgResearch, Ruakura Campus.

International Conference on Statistics, Combinatorics and Related Areas. *December 19-21, 2001, University of Wollongong, Australia.*

The School of Mathematics and Applied Statistics will be hosting this annual conference. Symposia on a number of topics are being organised. Some of the topics are: cointegration, computer security, data mining, ecological and/or environmental statistics, estimating function and quasi-likelihood, etc. (for a complete listing, check the web site below). Abstracts in plain-text are due by September 15, 2001. Further information at the web page: `www.uow.edu.au/informatics/maths/statconference/`

2nd MaPhySto Conference on Lévy Processes - Theory and Applications. *January 21-25, 2002, University of Aarhus, Denmark.*

This Conference is a sequel to the international conference, with the same title, organized by MaPhySto 7-22 January 1999. As a result of the interest engendered by the first conference, it was decided to have Lévy processes as an ongoing thematic topic for MaPhySto activities, for several years. Information on past such activities can be found at the MaPhySto website. Deadline for the registration is December 1, 2001. A list of invited speakers with further information regarding the conference may be found at the web site: `www.maphysto.dk/events/2ndLevyConf2002/`

✻ **I COBAL**. *February 3-7, 2002, Ubatuba, Sao Paulo, Brazil*.

The Brazilian Section of ISBA, co-sponsored by ISBA and the

Brazilian Statistical Association (ABE), is organizing the First Latin American Meeting on Bayesian Statistics (I COBAL). For more details see `www.est.ufmg.br/cobal`.

ProbaStat 2002. Fourth International Conference on Mathematical Statistics. *February 4-8, 2002, Smolenice Castle, Slovak Republic*.

The 4th international conference ProbaStat 2002 will be held in Smolenice, Slovakia. Contributions are supposed to cover some of the following areas: Regression models, Time series models, Experimental design, Applications. Invited talks on these subjects are expected as well as contributed talks and/or posters of the participants. Deadline for registration is September 30, 2001. Deadline for payments and submission of Abstracts is November 18, 2001. Preliminary conference information is available at `www.um.savba.sk/lab_15/probastat.html`.

Sixth International Conference on the Statistical Analysis of Textual Data . *March 13-15, 2002, St-Malo, France.*

The International Conference on the Statistical Analysis of Textual Data provides a workshop-style forum to all scholars, statisticians, computer scientists, linguists, etc., working in the vast field of textual data analysis ranging from lexicography to the analysis of political discourse, from documentary research to marketing research, from computational linguistics to sociolinguistics, from the processing of data to content analysis. Participants wishing to submit a paper should send to the organization committee (via email) a first version of their paper for review by September 1, 2001. Web page of the conference: `www.irisa.fr/manifestations/2002/JADT/welcome.htm`

Third International Conference on Mathematical Methods in Reliability. *June 17-20, 2002, Trondheim, Norway.*

The MMR 2002 is the third in a series of conferences on Mathematical Methods in Reliability. The MMR conferences intend to serve as a forum for discussing fundamental issues on mathematical methods in reliability theory and its applications. The idea is to assemble researchers from university laboratories and research institutions of probability, statistics and applied mathematics working in the reliability field and related fields, e.g. survival analysis. The program will include invited talks, plenary sessions, parallel sessions and posters on a broad range of topics. The presentations may range from theoretical studies to industrial applications. Deadline for submission of abstracts is October 15, 2001. See the Conference web page `www.math.ntnu.no/mmr2002/`

8th Vilnius Conference on Probability Theory and Mathematical Statistics. *June 23-29, 2002, Vilnius, Lithuania.*

The Institute of Mathematics and Informatics and the Vilnius Gediminas Technical University organize the 8th Vilnius Conference which will consist of a few plenary talks and sessions on important topics in Probability Theory and Mathematical Statistics. Prof. Vytautas Statulevicius is the Chair of the Organizing Committee. E-mail address for contact information is `conf@ktl.mii.lt`. Home page of the conference, where the interested reader may find more details, is: `www.science.mii.lt/vilconf8/`

Sixth International Conference on Teaching Statistics (ICOTS-6). *July 7-12, 2002, Durban, South Africa.*

The conference is organized by the International Association for Statistical Education and by the International Statistical Institute. The major aim of ICOTS-6 is to provide the opportunity for people from around the world who are involved in statistics education to exchange ideas and experiences, to discuss the latest development in teaching statistics and to expand their network of statistical educators. The conference will include keynote speakers, invited speakers, contributed papers, workshops and forums, demonstration lessons, roundtable sessions, poster sessions, book and software displays, hands-on computer sessions and many opportunities for the communication and exchange of experiences and ideas. As the conference theme for ICOTS-6 is

"Developing a statistically literate society", special sessions on statistics literacy are planned. Deadline for submission of contributed papers to be refereed is October 1, 2001. Web page of the conference: `www.beeri.org.il/icots6/`

First Madrid Conference on Queueing Theory. *July 15-21, 2002, Complutense University of Madrid, Spain.*

The Madrid Conference on Queueing Theory (MCQT) aims to be a meeting place where scientists and technicians in the field could find a discussion forum to promote research, encourage interaction and exchange ideas. The conference is open to all trends including the development of the theory, methodology and applications of Queueing Theory. Deadline for paper submission: September 1, 2001. A selection of papers presented at the conference will be published in a special volume of Annals of Operations Research (AOR). Further details may be found at he web page: `www.mat.ucm.es/deptos/es/mcqt/conf.html`

➤ Internet Resources

StatCalc 1.1

The StatCalc software is like a PC calculator that computes table values and other statistics for 34 probability distributions. A helpfile is provided for each distribution. It also includes some nonparametric table values, tolerance factors, and bivariate normal distribution. The software is free, and can be downloaded from the web site `www.ucs.louisiana.edu/~kxk4695/`

Software for Distance Sampling.

Distance is a Windows-based computer package that allows for design and analysis of distance sampling surveys of wildlife populations. The latest version of Distance is 4.0, which is currently in public beta. The version on general release is Distance 3.5. Web site for download is `www.ruwpa.st-and.ac.uk/distance/distanceabout.html`

MX Software

Mx is a matrix algebra interpreter and numerical optimizer for structural equation modeling and other types of statistical modeling of data. It enables exploration of matrix algebra through a variety of operations and functions. It offers facilities for maximum likelihood estimation of parameters from missing data structures, under normal theory. Complex 'nonstandard' models are easy to specify. For further general applicability, it allows the user to define their own fit functions, and optimization may be performed subject to linear and nonlinear equality or boundary constraints. Software can be downloaded at `views.vcu.edu/mx/executables.html`

Bayesian Inference for Physical Sciences.

BIPS (Bayesian Inference for the Physical Sciences) is an annotated online index/clearinghouse for information on the Bayesian approach to statistical inference of special relevance to applications in the physical sciences. It contains sections with information concerning General Texts and Tutorials on Bayesian Statistics, Preprints/Reprints, Bayesian Software and general statistical resources with Bayesian content. The web site address is `astrosun.tn.cornell.edu/staff/loredo/bayes/`.

MacTutor History of Mathematics Archive.

Prof. E. Robertson and Dr. J. O'Connor, from the University of St. Andrews, have been developing an excellent web site dealing with the history of Mathematics. It is an integrated collection of over 1000 biographies and historical articles of a mathematical nature, alongside interactive birthplace maps and the famous curve applet. Therein one can also find references to the biographies of many statisticians. The address is `www-history.mcs.st-and.ac.uk/history/`

Queueing Theory Page.

The Queueing Theory web site contains information on Queueing Theory collected by Dr. Myron Hlynka, of the University of Windsor, in Windsor, Ontario, Canada. It includes a list of books on queueing theory, a list of home pages and a list of e-mail addresses of queueing theorists. Web page address is `socr.uwindsor.ca/~hlynka/queue.html`

➤ Awards and Prizes

Research Prize Lectureships

The Canadian Mathematical Society Research Committee invites nominations for three prize lectureships

– The Coxeter-James Prize Lectureship recognizes outstanding young research mathematicians in Canada. The selected candidate will deliver the prize lecture at the Winter 2001 Meeting in Toronto, Ontario.

– The Jeffery-Williams Prize Lectureship recognizes outstanding leaders in mathematics in a Canadian context. The prize lecture will be delivered at the Summer 2002 Meeting in Quebec, Quebec.

– The Krieger-Nelson Prize Lectureship recognizes outstanding female mathematicians. The prize lecture will be delivered at the Summer 2002 Meeting in Quebec, Quebec.

The deadline for nominations is September 1, 2001. Letters of nomination should be sent to: Douglas Stinson, CMS Research Committee Department of Pure Mathematics, University of Waterloo 200 University Ave West, Waterloo,ON Canada N2L 3G1. For further information visit the web site:
`www.cms.math.ca/Prizes/`

Alice T. Schafer Mathematics Prize.

The Executive Committee of the Association for Women in Mathematics calls for nominations for the Alice T. Schafer Mathematics Prize to be awarded to an undergraduate woman for excellence in mathematics. All members of the mathematical community are invited to submit nominations for the Prize. The nominee may be at any level in her undergraduate career. Nominees must either be a U.S. citizen or have a school address in the U.S. Schafer Prize Nomination deadline is October 1, 2001. Detailed information are available at the following address:
`www.awm-math.org/schaferprize.html`

➤ Miscellanea

Bayesian seminar in Argentina.

On June 9th, the first Seminar on Bayesian Statistics was held under the joint organization of Universidad Nacional de Quilmes and Universidad Nacional de la Matanza.

The first part was an invited paper by Prof. Aldo Sacerdoti (`sacerdo@unlm.edu.ar`), about his experience in teaching Bayesian Statistics in undergraduate courses at Universidad de la Matanza and Universidad de Buenos Aires. He and his staff are also teaching the frequentist approach compared with Bayesian one. After his exposition, there was a discussion moderated by Prof. Alfredo Russo (Universidad Nacional de Quilmes).

The second part was a dissertation by Prof. Alfredo Russo about the works of his group about MCMC applied to project evaluation and monitoring with some examples extracted from their field work. A final discussion was moderated by Prof. Aldo Sacerdoti (Universidad Nacional de La Matanza).

Mailing address: *ISBA Bulletin - CNR IAMI - Via Ampère 56 - 20131 Milano (Italy)*
E-mail: *isba@iami.mi.cnr.it* **Phone:** *+39 0270643206* **Fax:** *+39 0270643212*
Web page: *www.iami.mi.cnr.it/isba*