THE ISBA BULLETIN

Vol. 8 No. 1 *March 2001*

The official bulletin of the International Society for Bayesian Analysis

A WORD FROM THE PRESIDENT

by Alicia Carriquiry
ISBA President
alicia@iastate.edu

As I write this first report as President of ISBA, the Executive Committee is preparing for its first annual meeting. Next week, the ISBA Executive Committee will be meeting in person at a gathering other than an ISBA world meeting or a Valencia meeting. This event is reflective of the maturity that our society has reached in its eleventh year of operation.

But before I look forward and tell you about the agenda for the upcoming meeting, I'd like to thank, on behalf of the membership, all of the ISBA officers who have recently completed their terms in office. These include John Geweke and Robert Wolpert, Past President and Past Chair of the ISBA Program Council in 2000, respectively, Mike Evans, former Secretary of ISBA (now acting as our very capable Web Master), and past members of the Board Alan Gelfand, Jay Kadane, Rob Kass, and Luis Pericchi. We welcome two new members of the Executive Committee, David Draper (President Elect) and Cindy Christiansen (Executive Secretary), a new member of the Program Council, Luis Raul Pericchi, and four new members of the Board of Directors: Nicky Best, Eduardo Gutiérrez Peña, Tony O'Hagan, and Raquel Prado.

After the successful world meeting in Crete, we are now well positioned to take advantage of new opportunities and to face new challenges. How can we best advance our mission of promoting the development of Bayesian theory and methods, and their application to problems in other disciplines? Items for discussion at the meeting of the Executive Committee will move us one step closer to fully meeting our goals as a society.

Recruiting and retaining new members is, as always, a top priority. In spite of the explosion in the use of Bayesian methods in most areas of application, and in the increasing number of our colleagues who conduct research in Bayesian statistics, our membership numbers have remained relatively flat, with the occasional spike following the world meetings. An effective way to attract additional members to ISBA is to forge closer links to other societies. In this regard, a promising development is the planning of an international conference, to be jointly sponsored by ISBA and the IMS, and that will take place possibly in the year 2003. Tony O'Hagan, Chair of the Program Council, will be reporting on this initiative in a later issue of the Bulletin.

ISBA has now the opportunity to co-sponsor the very successful Valencia meetings, starting with Valencia VII, to be held in 2002. The Organizing Committee of the Valencia International Meetings has proposed that ISBA plays a major role in setting the scientific program for Valencia VII; establishing a mechanism for soliciting and selecting contributions that is inclusive

Contents

and efficient is one of the items in the agenda for the meeting of the Executive Committee. This is a new milestone for ISBA, and one that will result in added benefits for our members.
Endowments for three awards are now administered by ISBA on behalf of the Bayesian community. In addition to the Savage Award endowment, that was transferred to ISBA for its administration last year, two new awards will be also be under the management of ISBA. The DeGroot Prize will be awarded every other year, starting in 2002, to the author(s) of a published book in statistical science. The prize honors the memory of Morris DeGroot, his contributions to statistics and decision theory, and in particular, the value that he placed on the role of books. The Lindley Prize will be awarded in 2003 and every other year after that, and is named after Dennis V. Lindley. Authors of innovative research that is accepted for publication in the refereed proceedings of the Valencia and ISBA international meetings will be eligible for the award. Both the DeGroot and the Lindley Prizes were generously endowed by an initial financial contribution from a large group of academic and other institutions. The charters with details on the administration, review process, and selection of winners for each of the awards can be found in the ISBA web site at `www.bayesian.org`.
As we think of expanded outlets and intellectual opportunities to provide to our members, it is inevitable that we revisit the possibility of publishing a journal in statistics. Whether to publish a journal or not, and if so, whether it ought to be published in electronic or hard paper format, by ourselves or in cooperation with other societies, and whether the journal should have a broad statistical scope, are some of the questions posed to an ad-hoc committee established last year by Past President Phil Dawid. The ad-hoc committee, composed of Rob Kass, Sylvia Richardson, David Heckerman, John Geweke, and me, will report within the next few months to the ISBA Executive Committee and to the Board of Directors. Because this is such an important topic, with potentially significant impacts for our society, the Bayesian community, and the profession in general, we will solicit input from the membership and will provide ample opportunity for discussion. In the past, our membership has been in (sometimes spirited) disagreement on the "journal question". Now, with a strong and definitely mainstream Bayesian community, it may be a good time to reconsider the pros and the cons of a possible ISBA publication to complement the revitalized ISBA Bulletin.
The Proceedings of ISBA 2000 will appear later this year. Edward George, the Editor for the volume, led the effort to solicit and gather contributions in a timely and efficient manner, and to oversee the reviewing of the manuscripts that were submitted for consideration. A selection of those papers that relate to official or government statistics and that were presented at ISBA 2000 will also appear in a special issue of the *Research in Official Statistics*, an international journal published by Eurostat.
I end my first "official address" by encouraging all of you to actively participate in our society. I know I speak for the rest of my colleagues in the Executive and in the Board when I say that we welcome your feedback, suggestions, comments and criticisms. ISBA is, for the first time, in good financial shape, and we now have an opportunity to think about creative ways in which we can better serve our membership. I look forward to an exciting year, and to providing stewardship for ISBA and our community in the months ahead.

BULLETIN BY E-MAIL

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

OUTGOING PRESIDENT'S REPORT FOR THE YEAR 2000

by Philip Dawid
Past ISBA President
dawid@stats.ucl.ac.uk

The millennial year 2000 was an active one for ISBA. A number of end-of-year reports from officers and committees, as well as minutes of Board and General Meetings, are now or will soon be made public on the ISBA website (www.bayesian.org), so I shall just mention a few highlights here.

The most notable ISBA activity was without doubt the highly successful ISBA 2000 meeting in Crete. Many people put much effort into organising this, with results that were greatly appreciated by all participants. Thanks to the generosity of Eurostat, all ISBA members will be receiving a free copy of the Proceedings in due course. Further details can be found in the reports of the ISBA Program Council, the ISBA 2000 Scientific Committee, and the ISBA 2000 Proceedings Editor.

Other noteworthy developments relate to prizes. At the request of the trustees of the Leonard J. Savage Memorial Fund, responsibility for managing the Savage Award for the best Bayesian doctoral dissertation (recently increased to two annual awards of $750 each, one for Theory and Methods, and one for Applied Methodology) has been transferred to ISBA. The ISBA by-laws have been augmented to reflect this. We have established a Savage Fund Committee, consisting of Ehsanolah Soofi (Chair), Peter Rossi, Arnold Zellner, Wes Johnson, Michael Evans and Clare Marshall, which includes representatives of the award's co-sponsors: ASA Section on Bayesian Statistical Science, and NBER-NSF Seminar on Bayesian Inference in Econometrics. In addition, the retiring Savage Fund trustees (Stephen Fienberg, Seymour Geisser, Ed Leamer, Jay Kadane and John Pratt) are emeritus members of this committee. A call for nominations for the 2000 awards has been made and is now closed. The evaluation committee will be chaired by Siddhartha Chib. Further details may be found on the ISBA website under "News" and "ISBA Awards".

At the ISBA 2000 meeting the organisers of the Valencia conferences decided to initiate two new prizes, to honour towering figures in Bayesian Statistics. The Dennis V. Lindley prize will be awarded biennially for the best contributed paper in the refereed proceedings of the Valencia Meeting or of the ISBA World Meeting. The Morris H. DeGroot prize will be awarded biennially to the author or authors of a published book in statistical science. Capital funding for these prizes has been obtained from a number of founders, including ISBA. Again, responsibility for managing these prizes has been transferred to ISBA, and further details may be viewed on the website.

Other business conducted by ISBA during the year included some revision of the by-laws regarding membership and dues, and the constitution of committees to consider the possibility of electronic publishing of a journal, and to increase the usefulness and informativeness of the website for interested non-members. On 16 March 2000 ISBA was (re)incorporated in the State of Iowa by Alicia Carriquiry.

During the course of the year a new Brazilian Chapter of ISBA was formed, to join existing chapters in India, South Africa and Chile. See the website under "Local Chapters". The ISBA Newsletter has been "relaunched" as the ISBA Bulletin, and goes from strength to strength under the energetic editorship of Fabrizio Ruggeri.

You will have noticed that ISBA has a smart new logo. This is Petros Dellaportas's winning entry in our logo competition. Thank you Petros, and I hope you enjoy your prize – a free holiday back at the ISBA 2000 hotel venue in Crete, thanks to the generosity of the hotel manager. Thanks too to all who voted, thus providing valuable testing of the new system for electronic voting which has since been used for the ISBA elections.

As I leave the Oval Office, I should like to express my sincere thanks (and distribute presidential pardons) to all those who are also stepping down at this time from service to ISBA. John Geweke has completed his term as Past President, and retires from the Executive and the Board. Robert Wolpert, having served as Program Chair in 1999, now rotates off Program Council.

Michael Evans relinquishes his position as Executive Secretary, although he will be continuing as Webmaster. The retiring Board members are Alan Gelfand, Jay Kadane, Rob Kass and Luis Pericchi, who served as ISBA Directors from 1997 to 2000. All these individuals have contributed greatly to the organisation and smooth running of ISBA, for which they deserve your gratitude as well as mine.

It only remains to congratulate all those who succeeded in the recent ISBA elections, and to wish Alicia and her team all success for the coming year.

A WORD FROM THE EDITOR

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

I em entering my last year as the Editor of the Bulletin and it is time to think about the future. It is time for all the ISBA members to give advice to the ISBA officers, Alicia Carriquiry *in primis*, on which directions the Bulletin should take under a different Editorial Board. The ISBA Executive Committee will decide, in a due time, how to proceed in the appointment of a new Editor. In the meanwhile, I urge you to let ISBA know what you think of the organisation of the journal and the way it is delivered.

We started with some sections we thought could be interesting for the readers. Over the last two years you have been able to see what we published and, may be, notice differences among past and new Associate Editors in the choice of the papers. You have enough evidence to comment on what you would like to keep and what you think is worth a change or needs to be discontinued. The new Editor, still to be appointed, will be helped by your valuable comments.

As I made it clear when I accepted the appointment as the Editor of a Newsletter which needed a strong intervention to revitalise it, I was not going to focus on the appearance of the Newsletter but mainly on its content. Now the Bulletin is doing fairly well, although some changes are welcome, and I believe it is time that more efforts will be devoted to its graphical aspect. The current Bulletin needs a lot of "cut and paste" in assembling the papers and uses no fancy software (just plain LaTeX).

I have been trying to convince ISBA members to ask for an electronic version of the Bulletin instead of the printed one. I am in favour of an electronic-only version to reduce costs and increase timeliness in the delivery. Your opinions on the subject are welcome.

Finally, I wish to warmly welcome all the Associate and Corresponding Editors who are starting their cooperation with the Bulletin with the current issue.

P.S. Some of you will have already noticed the absence of the usual interview: I owe all of you an apology. We have been working hard to get an interview for the current issue. We contacted, sequentially, two very famous statisticians but, regretfully, both of them were unable to answer our questions. They have been so kind with us, they tried hard to get the interview done but compelling reasons did not make it possible.

ISBA/SBSS ARCHIVE FOR ABSTRACTS

All authors of statistics papers and speakers giving conference presentations with substantial Bayesian content should consider submitting an abstract of the paper or talk to the ISBA/SBSS Bayesian Abstract Archive. Links to e-prints are encouraged. To submit an abstract, or to search existing abstracts by author, title, or keywords, follow the instructions at the abstract's web site,

www.isds.duke.edu/isba-sbss/

JIMMIE SAVAGE: AN INTRODUCTION

by Jay Kadane
kadane@stat.cmu.edu

Leonard J. ("Jimmie") Savage, was born in 1917. He was trained in mathematics at Michigan, and spent a postdoctoral year at Princeton with John von Neumann in 1941-42. Through his war-time work at the Statistical Research Group he moved into statistics, and by 1949 was a member of the Chicago Statistics Department, where he served as chair from 1957–1960. After a brief stay in Michigan, he moved to Yale in 1964 where he taught as Sterling Professor until his death in 1971 at the age of 54.

Savage came of intellectual age at an exciting time in the development of statistics. The book of von Neumann and Morgenstern (1944), building on earlier work of von Neumann (1928), explored the theory of games and proposed the minimax solution for two-person zero-sum games. They also gave axioms of preference which yielded maximization of expected utility as the correct way to make decisions. (Their probabilities were "objective," however). Following on the lead, Wald (1950) proposed using the minimax idea as a way to rationalize statistics, thinking of the statistician as playing a zero-sum two-person game against "nature."

Savage's (1951) essay in review of Wald's (1950) is respectful, but also points to difficulties in the minimax approach. He cites work of deFinetti, but only to suggest that the minimax approach might be suitable to group decisions.

Savage wrote his most important book (1954) to develop axioms to justify the minimax approach to statistics. The first seven chapters of his book state axioms for probability and utility jointly that lead to the maximization of subjective expected utility. The last ten chapters seek to apply this theory to justify frequentist practice. This latter half of his book is not very successful, of course, since the Bayesian foundations do not justify frequentist procedures. Savage himself (1970) later described his position in 1954 as "though interested in personal probability, ... , not yet a personalistic Bayesian and ... unaware of the likelihood principle."

Only gradually did Savage come to appreciate the importance and implications of his theory. By 1962, he reports "Though Savage (1954) emphasizes the merits of the concept of subjective (or personal) probability, it was not written in the anticipation of radical changes in statistical practice. The idea was, rather, that subjective probability would lead to a better justification of statistics as it was then taught and practiced, without having any urgent practical consequences. However, it has since become more and more clear that the concept of subjective probability is capable of suggesting and unifying important advances in statistical practice" (Savage, 1962, p. 9). By this time he was well aware of the likelihood principle, and more strongly connects his ideas to deFinetti's.

In addition to his training in mathematics, Savage's most intense applied collaborations were in economics, where he worked with and acknowledges some of the most important figures of his time, notably Milton Freidman. He also worked with biologists, especially just after World War II. But his published work does not reflect his wide-ranging curiosity about nearly every empirical matter. He supervised statistics Ph.D. students doing applied work in a wide variety of areas.

Savage was a controversial figure at the time, perhaps most importantly because he was advocating a position about the foundations of statistics that was not operational, in the sense that it was not possible to do the analyses and carry out the calculations suggested by his foundational position. This left statisticians with several unpalatable choices: (1) stop doing applied statistics, because doing it "right" wasn't possible (2) object to Savage's argument on some specious ground, in the hope of restoring respectability to statistical arguments one suspects aren't, in the end, all that sound (3) keep using the old methods, admitting their foundational weaknesses, resulting in a less confident approach of statisticians to their applied collaborators and clients, or (4) get angry at Savage for putting them in such an uncomfortable situation.

(As Dennis Lindley points out to me, frequentist statisticians also had grave difficulty at the time doing the calculations their theory suggested).

Savage was aware of the dilemma he created for statisticians, and was moving to address these issues in the last decade of his life. In his 1962 "discussion" he was soliciting statisticians to help in the work of making his ideas operational. He appealed to psychologists in Edwards, Lindman and Savage (1963). He gave his ideas about the critical issue of elicitation in his (1971) paper, and he gave his views on the crucial work of Fisher in his (1976).

A reader of Savage's papers can enjoy, in addition to the profundity of the ideas in them, the wonderfully clear style in which they are expressed. He cared about grammar, clear notation, mathematical simplicity, succinct examples, and precise expression.

This sketch would be incomplete without recounting my interactions with him. My first job, after graduating from Stanford in 1966, was at Yale, in the Department of Statistics whose Professors were Jimmie Savage and Frank Anscombe. My thesis and other work were entirely classical; I had heard a bit about Bayesian ideas, but had been told to concentrate on technique and leave matters of philosophical interpretation 'til later.

Jimmie was a wonderful mentor and colleague. He had time and patience for innumerable questions. He helped me a lot with my writing, which was so bad in those days that even I knew it was bad. I was writing a paper in search theory; Jimmie read it and marked it up 15 times, exposing flaws in logic, notation and expression each time. No-one before or since has given me that kind of attention. He even agreed to teach jointly with me from his book with Dubins (1965), which was the only way I saw to penetrate a very interesting, but very technical treatise. I stayed at Yale until the summer of 1968, but Savage was on leave in Italy (working with deFinetti) the spring semester of 1968. So it was only 3 semesters and a summer that I was with Savage. But that short exposure changed my view about statistics and posed the problems I have spent the rest of my professional career wrestling with.

An interested reader can find out more in no better way than by reading Savage. The places to start are with his masterpiece (Savage (1954), available in Dover Paperback) and a book of his papers together with several excellent summaries of his life and work, "The Writings of Leonard Jimmie Savage: A Memorial Selection" published by the American Statistical Association and the Institute of Mathematical Statistics, and available from ASA for the bargain price of $27.00.

➤ References

DUBINS, L.E. and SAVAGE, L.J. (1965). *How to Gamble If you Must: Inequalities for Stochastic Processes*. McGraw Hill.

EDWARDS, W., LINDMAN, H. and SAVAGE, L.J. (1963). Bayesian Statistical Inference for Psychological Research. *Psychological Review*, **70**, 193-242.

KADANE, J.B. (1968). Discrete Search and the Neyman-Pearson Lemma. *The Journal of Mathematical Analysis and Applications*, **22**, 156-71.

SAVAGE, L.J. (1951). The Theory of Statistical Decision. *Journal of the American Statistical Association*, **46**, 55-67.

SAVAGE, L.J. (1970). Reading Suggestions for the Foundations of Statistics. *American Statistician*, **24**, 23-27.

SAVAGE, L.J. (1971). Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association*, **66**, 783-801.

SAVAGE, L.J. (1976). On rereading R.A. Fisher (with discussion) (J.W. Pratt, ed.). *Annals of Statistics*, **4**, 441-500.

SAVAGE, L.J. ET AL. (1962). *The Foundations of Statistical Inference*. Methuen & Co.

VON NEUMANN, J. (1928). Zur Theorie der Gesellschaftsspiele. *Annalen Mathematische*, **100**, 295-320.

VON NEUMANN, J. and MORGENSTERN, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

WALD, A. (1950). *The Foundations of Statistics*. Wiley.

MAKING THE 21'st CENTURY BAYESIAN: TEACHING BAYESIAN STATISTICS AT FIRST YEAR LEVEL

by W. M. (Bill) Bolstad
bolstad@waikato.ac.nz

In 1975 D. V. Lindley predicted that in the 21'st century, statistics would be Bayesian. This appeared to be a very brave statement at that time, when the frequentist paradigm was firmly entrenched and the known theoretical advantages of Bayesian statistics were realizable in only the most simple cases. Events have proved Lindley to have had remarkable prescience. There were powerful Bayesian tools such as the Kalman filter, and the Metropolis algorithm which had been invented in other fields, but had yet to make much impact in statistics. During the past quarter of the 20'th century the use of these tools and others such as the Gibbs Sampler spread into statistics. Table 1 shows the percent of JASA articles clearly Bayesian by their title over the past 30 years.

Years	JASA Bayes articles
1971-1975	4%
1976-1980	5%
1981-1985	5%
1986-1990	6%
1991-1995	11%
1996-2000	15%

We can see clearly that the movement is in the right direction. However it is not moving nearly as fast as it should. When you have a better product, be it statistical paradigms or mousetraps, it is not satisfactory to see your competitor have such market dominance. Nevertheless, this table shows we have to realize that a majority of scientists are still comfortable with many of the ideas contained in frequentist statistics, despite their shortcomings. Evaluating procedures by their long run behavior is perceived as sensible by most scientists. Scientists use hypothesis tests as a tool for Occam's razor; it keeps them from going to overly elaborate models when a more simple one suffices. While not optimal because it doesn't take the consequences of the decision into account, it suits their purposes.

We must reach the future scientists with our message *Bayesian Statistics can be used to achieve their aims better than frequentist statistics.* Most students only do one statistics course. This means we have an *Introduction to Bayesian Statistics* as the first course for our best students, or we will fail to reach the majority of potential scientists. If we get the future scientists thinking Bayesian from the start, they will be doing Bayesian analysis when they write up their research, and asking for it when refereeing other scientific papers. They will know the advantages of the Bayesian approach, which can now be achieved using MCMC methods. These methods, however, are beyond the level of the first course, and are introduced in subsequent courses.

The philosophy behind the first year *Introduction to Bayesian Statistics* course I have developed at the University of Waikato is that the topics covered will be the very similar to those in a first year *Introduction to Statistics* Course taught from a frequentist perspective. It is a course on Bayesian statistical inference, not decision theory. I think it is more important to introduce a few of the key ideas well, than introduce a coherent, though confusing welter of ideas to be sorted out. This leaves more time for the practical issues such as how to determine a prior, and judge its suitability.

I target the course at students who have either some understanding of calculus, or are concurrently enrolled in a calculus course, mainly to get students who are not formula averse. Using the rules of probability requires that much. Students must understand that an integral represents an area under the curve, but don't have to do any actual integrating analytically. They learn to recognize the part of the density that determines the shape, and put in the constant that makes it integrate to one.

The course starts by emphasizing the importance of data gathering design. This includes simple, stratified, and cluster random sampling, independent and paired experimental designs. We note the difference between observational studies, and randomized experiments, and the different conclusions that can be drawn from them. Real data sets are used whenever

possible.

Bayesian inference methods are developed for the same models as in a frequentist introductory course: (binomial) proportions, normal mean, difference between normal means, difference between proportions, and simple linear regression.

Subjective priors are elicited from the conjugate family by matching moments. The importance of graphing the prior to make sure that it reasonably reflects your belief is emphasized. The *equivalent sample size* is used to prevent students using a prior that is too precise relative to the sample size. Flat priors are used to represent prior ignorance. We demonstrate that using any reasonable prior doesn't change the posterior very much.

We use the posterior mean as an estimator. We show it has excellent characteristics when averaged over the sample space and performs better than the corresponding frequentist estimator in terms of mean square over the most of the possible parameter values. We introduce credible intervals for the parameter and contrast their useful interpretation with the backwards interpretation of the corresponding frequentist confidence interval. Hypothesis testing is well entrenched into science, so we feel we have to introduce it, but in a Bayesian manner. One sided tests are performed by rejecting the null hypothesis whenever the posterior probability of the null hypothesis is below the level of significance. We test the credibility of a point null hypothesis versus a two sided alternative by looking at whether the null value lies inside the corresponding credible interval.

We introduce the Student's t distribution as an approximation to be used when the variance is estimated form the sample. The idea of marginalization is introduced, but the calculations are beyond the level I wish to go in an introductory class. The only place where we do marginalize is when we find the predictive distribution of a new observation by marginalizing out the parameter.

We have a tutorial session each week where the students discuss and work through a statistical activity covering the previous weeks topic. I also introduce the idea of Monte Carlo studies. Statistical computing using Minitab is an essential part of the course. There are about 4 Minitab projects including some small scale Monte Carlo studies comparing long run behavior of the posterior Bayesian estimators with the frequentist alternative. I have written a Minitab macro which students can use to do any numerical integration required in the projects. I have taught this course several times at the University of Waikato over the past few years, always with good results. I am currently preparing a textbook based on the course. Details about my course can be obtained by emailing *bolstad@waikato.ac.nz*.

BAYESIANS IN TURKEY

by Gül Ergun
gul@hacettepe.edu.tr

Firstly, I would like to express my gratitude to the editorial board for giving me the chance of presenting Turkey in this occasion.

Bayesian statistics in Turkey is at its early stages. Therefore, it seems quite difficult to talk about existence of related studies and progress in this specific area. A historical view at the position of statistics may shed some light on the level of Bayesian Statistics in Turkey. The first department of statistics was established at Hacettepe University in 1967. Today, the number of departments of statistics have reached 16, most of which have been established within the last ten years. As seen from these figures, the history of education on statistics in Turkey is fairly new. This is one of the most important reasons for late tendency toward Bayesian statistics.

There is no Department of Statistics in Turkey which only deals with pure Bayesian approach. Unfortunately, this situation hinders the chance of gathering a specific research team that studies Bayesian Statistics.

Even a brief look at statistics community in Turkey may show that most of the statisticians are non-Bayesian. However, it seems quite interesting that

they accept the Bayesian approach as a useful method with a good theoretical base in statistical decision theory. The names and research areas of some statisticians who are not pure Bayesian statisticians but use Bayesian approach in their research are as follow:

• Ulkü Gurler (ulku@bilkent.edu.tr): Reliability Theory and Inventory Control
• Suleyman Ozekici (ozekici@boun.edu.tr): Reliability Theory
• Emre Berk (berk@bilkent.edu.tr): Inventory Control
• Omer Gebizlioglu (gebizli@science.ankara.edu.tr): Risk Theory and Decision
• Ceyhan Inal (cinal@hacettepe.edu.tr): Stochastic Processes
• Suleyman Gunay (sgunay@hacettepe.edu.tr): Hypothesis Testing and Decision
• Aydin Erar (erar@hacettepe.edu.tr): Bayesian Variable Selection in Regression
• Atilla Yardimci (atilla@yardimci.net): Bayesian Variable Selection in Regression
• Sevil Bacanli (sevil@hacettepe.edu.tr)
• Ozgur Peker (opeker@Anadolu.edu.tr)

In addition to these names Ihsan Alp (ialp@quark.fef.gazi.edu.tr) and Resat Kasap (rkasap@quark.fef.gazi.edu.tr) may be added to this list for the future Bayesian teams. However as it can be seen, the number of Bayesian statisticians in Turkey is not that high, there are also Turkish Bayesian Statisticians who are working abroad on this specific area, namely Refik Soyer, Alaatin Erkanli, Mehmet Sahinoglu. They still have ties in Turkey. These statisticians had active roles as boarding members on ISBA'97 World Meeting which was held in Istanbul.

I have been interested in Bayesian Statistics since 1990, the time I started my MSc studies in Warwick University. I was very impressed by the Bayesian statisticians, especially A. O'Hagan, J. Harrison and J.Q. Smith. Having adopted Bayesian philosophy, I decided to study Bayesian time series for my PhD as well. I have been lecturing at Hacettepe University, Department of Statistics since 1995. I am still working on Bayesian Time Series Modelling and MCMC techniques despite limited resources. I am trying to do my best to spread the Bayesian idea to my PhD, MSc and undergraduate students. However, it is quite difficult to spread the idea among people who have classical background.

I observe that Suleyman Ozekici, the former corresponding editor, also shares my thoughts that Bayesians in Turkey are limited in number. For this reason, I would also like to mention about the activities in the field of statistics in general. Turkish Statistical Association established in 1990 with the aims of guiding the scientific researches in Turkey, gathering the works of people both from the scientific and professional area, and representing Turkey on international platform. The Association currently has 500 members. The chair of the Association is Omer Gebizlioglu. The Association organized many national and international meetings and conferences. The next conference, entitled "Characterization Problems and Application", will be held on 1-3 May 2001 Antalya.

State Institute of Statistics also organizes national meetings every year on November. In addition to those, Statistics Graduates' Society organizes Symposiums every two years. The forthcoming mmeting of the society, 2^{nd} Statistics Symposium, will be held on 2-6 May 2001 in Antalya. Even though it is early to announce I would like to inform you that the international EURO/INFORMS meeting will be held in Istanbul in 2003. As a conclusion, the developments in Bayesian Statistics in Turkey seem at its early stages. However I am not pessimistic, I believe that the future is bright for Bayesian Statistics in Turkey.

BAYESIAN ANALYSIS OF A CLINICAL TRIAL

by George G. Woodworth
george-woodworth@uiowa.edu

The Iowa High Performance Driving Simulator, and its successor, the National Advanced Driving Simulator (www.nads-sc.uiowa.edu) provide a realistic, safe way to study the effects of drugs (Weiler, et al. 2000) and medical devices (Featherstone, et al 1999) on driving performance. These facilities are unique in providing a highly realistic driving simulation. This degree of realism is provided by wrap-around, real-time video simulation of the view through the windows of the vehicle. Simulation of the forces of acceleration and deceleration generated by braking, turning, avoidance maneuvers, etc. is provided by mounting the vehicle cab on a force platform driven by 6 large hydraulic actuators.

The study described here (Weiler, et al. 2000) was a four period double blind, double dummy crossover trial comparing the effects of fexofenadine (a non-sedating, prescription antihistamine), diphenhydramine (Benadryl$^{(r)}$), an over the counter, sedating antihistamine, alcohol, and placebo. In a crossover trial, each subject receives all treatments during successive periods. Forty subjects were randomized with respect to the 24 possible orders of administration of the treatments (ADFP, AFDP, etc.).

The treatment visits occurred weekly on the same day at the same time. Subjects were given two capsules (fexofenadine and diphenhydramine) 2.5 hours before each drive and were given a beverage 60 minutes before the drive. For the diphenhydramine and fexofenadine conditions, one capsule was active, the other was a placebo, and the beverage contained no alcohol (although alcohol was swabbed around the rim as a placebo). For the alcohol condition, both capsules were placebos and the beverage contained an amount of alcohol calculated to produce legal intoxication based on sex and body-weight. For the placebo condition, both capsules and the beverage were placebos.

The simulated drive was conducted in dry weather conditions with good visibility, on a two-lane rural highway that was 45 miles long. The lane widths were a standard 12 feet and the road surface was standard blacktop. The posted speed limit was 55 mph. Vehicles in the oncoming lane simulated low-density traffic. Participants practiced for 8 to 10 minutes before each experimental drive. For approximately the first 15 miles, the subject was instructed to follow a simulated lead vehicle maintaining a constant distance. The lead vehicle varied its speed smoothly but unpredictably. After the lead vehicle turned off the subject was instructed to drive "as you normally would".

Outcome measures included the subjects' accuracy in following the lead car, their steering instability, and the number of times any part of the vehicle crossed the center lane.

Steering instability is the root-mean-square deviation (in meters) of the participant's car around the participant's preferred position in the lane. Participants with high instability wandered left and right within (and sometimes out of) the lane. We measured deviations from the preferred position rather than the geometric center of the lane to avoid penalizing otherwise steady drivers who simply preferred to be closer to the centerline or to the right shoulder line. Not surprisingly, steering instability was highly correlated with center lane crossings.

Crossover designs have advantages and drawbacks. With four treatments, a crossover design requires one-fourth the number of participants required by a completely randomized design. Furthermore, because each participant acts as his or her own control, it is, in theory, possible to compare treatments with greater precision. The most problematic aspect of crossover designs may be the effect of previous experiences on a participant's reaction to the current treatment. Such effects can be broadly classified as period effects (learning, habituation), which are unrelated to previous treatments, and carryover effects, which are related to previous treatments. Although it is unlikely that any residual drug remained after an interval of 1 week, drug effects can carry over in other ways.

For example, if one of the drugs promoted simulator motion sickness, the participant

may have driven with more caution the week after receiving that drug. Another complication of crossover designs is the need to model the covariance structure of repeated measures on the same subject.

The data model included treatment, period and crossover effects,
$y_{sp} = \delta_{t(sp)} + \gamma_{t(s,p-1)} + \pi_p + e_{sp}$
where t(sp) is the treatment administered to subject s at period p (1=A, 2=D, 3=F, 4=P), $\delta_t,\ 1 \le t \le 4$ are the period-1 treatment means, $\gamma_t,\ 1 \le t \le 4$ are the lag-1 carryover effects (set to 0 for period 1), and $\pi_p,\ 1 \le p \le 4$ are the period effects. The residual vectors $(e_{s1}, \ldots, e_{s4})$ are independent between subjects and have 4-variate normal distributions with mean 0 and precision matrix T.

The prior distribution of the treatment means (δ_t), was i.i.d normal with mean μ and precision τ_δ. For identification, the carryover effect of placebo (γ_4) and the period 1 effect (π_1) were set to zero. The remaining carryover and period effects were given i.i.d. normal prior distributions with means 0 and precisions τ_γ and τ_π, respectively. The data precision, T was given a Wishart prior distribution with 4 degrees of freedom and identity scaling matrix. The grand mean, μ , had a normal distribution with mean zero and precision .01, and the precision parameters τ_δ, τ_γ, and τ_π had gamma distributions with means 1 and shapes .01.

The response variable, steering instability, had a strongly skewed distribution. Box-Cox analysis of the residuals (ignoring correlation) suggested the reciprocal transformation, which did produce a reasonably normal distribution of the residuals. Consequently, the final model was

$$y_{sp}^{-1} = \delta_{t(sp)} + \gamma_{t(s,p-1)} + \pi_p + e_{sp}$$

The model was implemented in WinBUGS version 1.2 and converged rapidly from a variety of initial values, showing little sensitivity to the values of the fixed hyperparameters.

The research team thought that readers might have trouble interpreting effect estimates reported on the transformed scale (meters-1), so the treatment effects were back-transformed to the meter scale, and contrasts were computed on that scale (for example, the D-P contrast is $\delta_2^{-1} - \delta_4^{-1}$). This is trivial for Bayesian analyses implemented via MCMC, and essentially impossible in a frequentist analysis. It was this impossibility that persuaded the statistical referee to accept a Bayesian analysis. The results are reproduced in Table 1.

	Posterior Distribution		
Parm	mean	$P_{2.5}$	$P_{97.5}$
A	.512	.498	.531
D	.527	.508	.546
F	.492	.477	.509
P	.495	.480	.513
A - D	-.014	-.029	.000
A - F	.020	.007	.033
A - P	.017	.003	.031
D - F	.034	.020	.049
D - P	.031	.017	.046
F - P	-.003	-.017	.010

Table 1. Posterior distributions of period 1 treatment effects and contrasts (meter scale).

Table 2, which may comfort the frequentist, shows that on the transformed scale, frequentist point estimates and confidence intervals (from SAS PROC MIXED) agree closely with posterior means and credible intervals.

		Inferential Statistics		
Parm		mean	$P_{2.5}$	$P_{97.5}$
A	B	2.048	1.983	2.116
	F	2.050	1.983	2.117
D	B	2.101	2.034	2.170
	F	2.105	2.038	2.171
F	B	1.969	1.905	2.035
	F	1.969	1.903	2.034
P	B	1.981	1.918	2.049
	F	1.981	1.916	2.047

Table 2. Inferential statistics for treatment effects on the (meter^{-1}) scale for Bayesian (posterior mean and credible interval) and Frequentist (maximum likelihood estimate, and asymptotic confidence interval) analyses.

➤ **References**

WEILER, J.M., WOODWORTH, G.G., GRANT, A.R, LAYTON, T.A., BROWN, T.L., MCKENZIE, D.R., BAKER, T.W. and WATSON, G.S. (2000). Effects of Fexofenadine, Diphenhydramine, and Alcohol on Driving Performance. *Ann. Intern. Med.*, **132**, 354-363.

FEATHERSTONE, K.A., BLOOMFIELD, J.R., LANG, A.J, MILLER-MEEKS, M.J, WOODWORTH, G.G. and STEINERT, R.F. (1999). Driving simulation study: Bilateral Array multifocal versus bilateral AMO monofocal intraocular lenses. *J. Cataract Refract Surg.*, **25**.

SOME WORDS ON THE R PROJECT

by Paulo J. Ribeiro Jr. and Patrick E. Brown
paulojus@est.ufpr.br
p.brown@lancaster.ac.uk

➤ Introduction

The choice of statistical computing environment is a recurring issue and is often a source of conflicting views. The decision of which programme(s) to adopt involves discussion of the needs of teaching, consultancy and/or research, and whether the software is for private use or for adoption at the workplace. The process can involve the intermingling of philosophies of software choice with constraints imposed by availability and finances. There is some consensus that availability of mathematical and statistical tools should be combined with an environment for implementation of new methods. Writing from the econometrics perspective Cribari-Neto and Zarkos (1999) point out that "computer work was mainly a question of reading the manuals and identifying which of the pre-packaged routines would perform the desired task. Times have changed, however, and many newly developed techniques are not available in econometrics packages". Their comments certainly apply to other areas which demand specific techniques or adaptations of the existing ones. From the statistician's perspective there will always be the need for a flexible environment to implement new methods and make them available to a wider audience.

In this article we describe the `R Project` and explain the main reasons why we adopted `R` as our preferred statistical computing environment. These are comments from the user's perspective, and we are not `R` specialists with in-depth knowledge about the inner workings of the system. Patrick has been using `R` since December 1999, judging by the date of his oldest `.RData` file, for both interactive use and computationally intensive batch jobs for his duties as a research associate. Paulo has been using `R` since January 2000 developing a package for geostatistical analysis as part of his PhD work and has run short courses with practical sessions using the package. The text is therefore biased towards `R`'s suitability for our specific research activities in an academic environment. It also reflects our status as computer-literate career statisticians who use the package on a daily basis.

The question we first asked ourselves when writing this article was: *why do we use R* ? The short answer was: it's a good quality statistics package that includes a slew of traditional and modern statistical methods. It's available for a number of operating systems, it's easy to create functions and packages for teaching and research, and the graphics are superb. It interfaces well to lower-level languages such as C, C++ and FORTRAN, and ... it's open source and free!

➤ Platforms

One of the original motivations for the development of `R` was to provide multi-purpose statistical software for the Linux operating system, and `R` was probably the first package to achieve this. Linux, because it is free and open source, has gained popularity not only amongst Unix users, but as an alternative to Windows. As Linux has evolved it has become easier to install and use, and there are currently a number of user-friendly interfaces that come bundled with most Linux distributions. As a consequence the numbers of universities, companies, and private users adopting Linux is increasing.

`R` is well suited to Linux. Many Linux distributions (Mandrake 7.2 for instance) come bundled with `R` (and Matlab's clone: Octave), as well as with many tools that `R` can exploit. `R` is also available for many Windows releases, Macintosh, and Solaris. Since the source code is distributed, `R` can be compiled to run with any platform. (Patrick is waiting for a PalmOS version!)

`R` is accessed solely via a command-line interface, even with the Windows version. This is because, as a non-commercial project, developers had other priorities. Besides, the debate about whether menu-driven interfaces cause more harm than good is still raging. Seasoned programmers will appreciate `R`'s terse interface, as it lends itself to speed and flexibility, but users of Excel and others menu driven software may find the

solitary ">" prompt intimidating. There is an attempt to develop a GNOME interface but this is still in its early stages. For Emacs users, the functionality of ESS (*Emacs Speaks Statistics*) includes support for R.

➤ So, what is R?

Quoting the R documentation: "R is a computer language not entirely unlike the S language developed at AT&T Bell Laboratories by Rick Becker, John Chambers and Allan Wilks. The two languages are implemented quite differently, but bear enough superficial resemblance that users should be able to switch between the two with relative ease." R can be regarded as a re-implementation of the S language. Motivations for its development include an attempt at a more efficient implementation of the S language, specially concerning memory demands, and development for platforms not supported by S-PLUS (at least at that time) like Linux and Macintosh. R was first announced on the *S-News* in August, 1993. Published reference about the project goes back to Ihaka and Gentleman (1996) reporting the initial development at the Statistics Department of the University of Auckland. Since 1995 R has been distributed as a free software with beta versions available and constantly updated. The first non-beta version (R-1.0) was released on 29^{th} February, 2000. The version 1.2 introduced significant changes in the memory management. The main consequence is the workspace is no longer static, it can grow and shrink as needed, freeing the user of the obligation to anticipate and allocate necessary memory for each session. The current version is R-1.2.1 available at CRAN (Comprehensive R Archive Network).

R consists of two basic parts: *R-base*, with the main code, and the *contributed packages*, a collection of codes implementing a diversity of statistical methods. In R's terminology the word *package* is used instead of S-PLUS's *library*. R development and policy are coordinated by the R *core team*, currently consisting of Douglas Bates, John Chambers, Peter Dalgaard, Robert Gentleman, Kurt Hornik, Ross Ihaka, Friedrich Leisch, Thomas Lumley, Martin Maechler, Guido Masarotto, Paul Murrell, Brian Ripley, Duncan Temple Lang, Luke Tierney. The core team was established in mid-1997 and plays the central role in R development and distribution, including responsibility for changes in the basic code. The project as a whole has contributions from a much larger group of people. The names of the main contributors are listed by typing `contributors()` at the R prompt.

For those of you unfamiliar with both, R and S-PLUS are packages which provide facilities for data manipulation, calculation and graphical display. These include collections of tools for data analysis and a simple programming language which can be used to implement new methods of data analysis and/or to customise existing ones. A key distinction from other statistical software is the emphasis on doing analyses in steps, storing results in objects which can be processed and interrogated by other functions. R is a run-time environment, meaning that code is not compiled into executables the way C code, for instance, is. A short example of R code is:

```
xyregression <- lm(y ~ x)
qqnorm(xyregression$resid)
printgraph(file="qqnorm.ps")
```

The first command performs a linear regression on the data in vectors x and y, saving the results as an object called "xyregression"; the second plots the normal scores of the residuals in a graphics window and the third saves the graph as a postscript file.

➤ R and S-PLUS: are they different?

The most obvious difference between R and S-PLUS is that R is an open-source project which distributes its source code under the terms of the GNU public license. The first advantage of open-source software is that it is free, freeing up a department's software license budget to fund more conferences overseas for postgrads and researchers! In fact, free software might be the only viable alternative for several departments around the world which cannot afford annual licenses for good quality statistical software.

On a scientific note, open source software allows the user

to inspect and study the source code, so we'll always know exactly which eigenvalue algorithm is used to compute our principle component analyses. Furthermore, the source code can be modified and re-distributed, or compiled as a stand-alone executable, provided the new code is also made publicly available.

The primary differences between the two packages are not on the surface, but on how the programmes function internally. Although to a user the programmes are almost identical, the bones of the packages are quite different. The root of the difference is that R uses what clever people call *lexical scoping*. This feature is inherited from the *Scheme* programming language from which R developers have borrowed many ideas. S-PLUS only uses local and global variables, whereas R has a hierarchy of variables whereby a function defined within a function can see variables defined in the first function. One result of the differing implementations is that R manages its own memory, allowing loops to operate more quickly than S-PLUS, reducing the critical problem of exploding memory usage. Ripley (2001) point out that: "R is more tolerant of badly-written code which can make S-PLUS slow to crawl". We both have had programmes that use too much memory to be run in S-PLUS but run without problems in R ...

For the average user the two packages appear almost identical. Changing the `motif()` command to `X11()` is sufficient to run most simple S-PLUS code in R. One important difference is that while S-PLUS saves each object as a separate file, R saves the entire workspace as one file, the way Matlab does. Transferring data between R and S-PLUS is usually straightforward with the `dump` and `restore` commands. There is no analogue of S-PLUS's Trellis on R. On the other hand R's plots accept mathematical notation and have more versatile colors schemes, at least when using the command line under Linux. Yet another difference is the extra flexibility in algorithms for random number generation: there are more algorithms available in R (and being open source they can be re-implemented in other engines). Other differences are listed in R's FAQ and Venables and Ripley's (1999) on-line complements. We will leave them to the reader. Compatibility with S has been pursued but sometimes has to be sacrificed in attempts to repair short-comings of S.

We particularly like the tools to create packages available to the R programmer. Documentation can be written using a, LaTeX style, pre-formatted type of document. A host of perl scripts assemble the code, check the example functions, build and convert documentation to a number of formats (*pdf* and *HTML*, among others). Scripts are available to convert to and from S-PLUS formats. Quite often we find packages originally written for one of the engines and converted to the other. Building a package like geoR say, under Linux, is done by: writing functions and documentation, copying them and other source code to a directory geoR with a specific structure, writing a short DESCRIPTION file and typing the shell command

```
R CMD build geoR.
```

The resulting file can be distributed and installed by typing:

```
R INSTALL
geoR_version.number.tar.gz.
```

For packages available at CRAN the installation and update is even easier. It can be done online during an R session using the functions `install.packages()` and `update.packages()`.

➤ Teaching

R is ideal for teaching for a number of reasons. Since it is accessed by programming rather than menus, students can perform linear regression by typing:

```
betahat <- solve(t(x) %*% x)
%*% t(x) %*% y
```

and can be stimulated to write their own code, as well as to inspect the ones already available.

Being free, R can be distributed and installed on personal computers. Class examples can be reproduced at home regardless of which engine is used at the University Labs. Packages can be easily prepared providing excellent and well organized material for courses. Actually, much of the available resources in R comes from teaching material. A very nice example of teaching

introductory courses using R is given by Nolan and Speed (2000). An experience with a web-based interface is described by M.J. Ray in the first issue of the R News. The article describes an implementation of a *Rcgi* interface at the University of East Anglia. Analyses can be performed remotely. Input is passed from a web-browser to the program running on a server and output is passed back again to the web-browser.

➤ **Integration**

As programming tools R and S are easier to learn and quicker to program though not as efficient as languages like C, C++ or Fortran. However, one of the virtues of both is the possibility of integration with other codes. Shared libraries can be loaded, opening the possibility of using R and S-PLUS for graphical interface even if the number-crunching uses a lower-level languages.

Other kinds of integration and interfaces have been discussed and implemented in R. Some examples are alternatives to distributed computing, interfaces with database management systems and communications with other languages/applications. Probably, these are areas where we will see most of the further developments. We refer to the first issue of the *R-News*

```
cran.r-project.org/doc/Rnews/
```

for more information on specific projects.

➤ **Bayesian methods in R**

Although the current distribution does not include many Bayesian tools, other available resources, including integration with other languages, make R a suitable environment for implementing Bayesian methods. The package CODA (output analysis and diagnostics for MCMC) by Martyn Plummer, Nicky Best, Kate Cowles and Karen Vines, included as a contributed package in CRAN, is an excellent tool for Bayesians needing to convince referees that their chains have indeed converged. Paulo has implemented some Bayesian methods for spatial data in the geoR package. There are certainly more Bayesian packages available that we will be chastised for not mentioning.

There have been messages on the topic of Bayesian methods on R's help list and it is clear that there is space for further development, either in the form of additional R packages or by integrating R with existing Bayesian software. Hopefully in the future there will be more resources at CRAN implementing Bayesian methods.

➤ **Documentation and Support**

The official R's web sites are:

```
www.r-project.org
cran.r-project.org
```

The first is "the" R home page while the second works as a download area with several mirrors. Extensive and detailed documentation can be found at the sites, and are also included in the software distributions. Some examples of available documents are ''An Introduction to R'' and ''Writing R Extensions''. Specific documentation for the packages can be found at the contributed packages web-page. Many basic questions are answered in R's FAQ. Functions are documented and help on specific functions are obtained by typing

```
help(function.name).
```

The function help.start() allows visualisation in HTML format. The only support available is through the *R help list* and we can only report positively from our experience subscribing to it.

➤ **How to start?**

The best way to start is to be motivated by an ongoing project. It's worth looking at the *demos* available by typing at the command line

```
demo()
```

and then selecting one of the available options. Try demo(graphics), for example. Users not familiar with S-PLUS can start running the examples listed in the Appendix A of ''An Introduction to R''. This will give a good idea of the R environment and the basic resources available, before start reading more detailed material.

Given the similarities with S-PLUS, most of the literature can be shared. For example the "classic" books by Becker, Chambers and Wilks (1998) and Chambers and Hastie (1988) can be adapted for R. Venables and Ripley (1999) has online complements on R and Venables and Ripley (2000) is written such that R's specific features are highlighted.

➤ **References**

BECKER, R.A., CHAMBERS, J.M. and WILKS, A.R. (1988) *The New S Language*. Chapman & Hall.

CHAMBERS, J.M. and HASTIE, T.J. (1992) *Statistical Models in S*. Chapman & Hall.

CRIBARI-NETO, F. and ZARKOS, S.G. (1999) R: yet another econometric programming environment. *Journal of Applied Econometrics*, **14**, 319–329.

IHAKA, R. and GENTLEMAN, R. (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–315.

NOLAN, D. and SPEED, T.D. (2000) *Stats Labs: Mathematical Statistics Through Application*. Springer.

RIPLEY, B.D. (2001) The R project in statistical computing. *LTSN Math, Stats & OR Network Newsletter* No 5 (March).

VENABLES, W.N. and RIPLEY, B.D. (1999) *Modern Applied Statistics with S-PLUS*. 3^{rd} ed. Springer.

VENABLES, W.N. and RIPLEY, B.D. (2000) *S Programming*. Springer.

BAYESIAN ANALYSIS OF CHOICE DATA

by Duncan K. H. Fong
i2v@psu.edu

We present recent Bayesian work on choice models which have many applications, in particular, in marketing.

Choice data are obtained when subjects are required to choose from a number of mutually exclusive alternatives. Partly because of the recent advance in computational algorithms like Markov Chain Monte Carlo methods, there is a growing interest in analyzing such data from a Bayesian perspective.

• J. H. ALBERT AND S. CHIB (1993). **Bayesian Analysis of Binary and Polychotomous Response Data**. *Journal of the American Statistical Association, 88, 669-679.*

In this article, exact Bayesian methods for modeling categorical response data are developed using the idea of data augmentation. This data augmentation approach provides a general framework for analyzing binary regression models. Under the proposed framework, the class of probit regression models can be enlarged by using mixtures of normal distributions to model the latent data. The methods can also be generalized to multinomial response models with $J > 2$ categories. In the ordered multinomial model, the J categories are ordered and a model is written linking the cumulative response probabilities with the linear regression structure. In the unordered multinomial model, the latent variables have a multivariate normal distribution with unknown variance-covariance matrix. For both multinomial models, the data augmentation method combined with Gibbs sampling is outlined.

• M. H. CHEN, D. K. DEY AND Q. M. SHAO (1999). **A new Skewed Link Model for Dichotomous Quantal Response Data.** *Journal of the American Statistical Association, 94, 1172-1186*.

In this article the authors propose a new skewed link model for analyzing binary response data with covariates. Introducing a skewed distribution for the underlying latent variable, they develop a class of asymmetric link models for binary response data. Using a Bayesian approach, they characterize the propriety of the posterior distributions using standard improper priors. They further propose informative priors using historical data from a similar previous study.

• R. MCCULLOCH AND P. E. ROSSI (1994). **An Exact Likelihood Analysis of the Multinomial Probit Model.** *Journal of Econometrics,* 64, 207-240.

The authors develop new methods for conducting a finite sample, likelihood-based analysis of the multinomial probit model. Using a variant of the Gibbs sampler, an algorithm is developed to draw from the exact posterior of the multinomial probit model with correlated errors. This approach avoids direct evaluation of the likelihood and, thus, avoids the problems associated with calculating choice probabilities which affect both the standard likelihood and method of simulated moments approaches. They also develop methods for analyzing random coefficient and multiperiod probit models.

• A. NOBILE (1998). **A Hybrid Markov Chain for the Bayesian Analysis of the Multinomial Probit Model**. *Statistics and Computing*, 8, 229-242.

Bayesian inference for the multinomial probit model, using the Gibbs sampler with data augmentation, has been considered by some authors. This paper introduces a modification of the sampling technique, by defining a hybrid Markov chain in which, after each Gibbs sampling cycle, a Metropolis step is carried out along a direction of constant likelihood. A proof of the ergodicity of the hybrid Markov chain is also given.

• P. E. ROSSI, R. E. MCCULLOCH AND G. M. ALLENBY (1996). **The Value of Purchase History Data in Target Marketing**. *Marketing Science*, 15, 321-340.

An important aspect of marketing practice is the targeting of consumer segments for differential promotional activity. The premise of this activity is that there exist distinct segments of homogeneous consumers who can be identified by readily available demographic information. A study is presented whose goal was to assess the information content of various information sets available for direct marketing purposes. Information on the consumer is obtained from the current and past purchase history as well as demographic characteristics. New econometric methods to implement a random coefficient choice model in which the heterogeneity distribution is related to observable demographics are developed.

• S. CHIB AND E. GREENBERG (1998).**Analysis of Multivariate Probit Models.** *Biometrika*, 85, 347-361.

This paper provides a practical simulation-based Bayesian and non-Bayesian analysis of correlated binary data using the multivariate probit model. The posterior distribution is simulated by Markov chain Monte Carlo methods and maximum likelihood estimates are obtained by a Monte Carlo version of the EM algorithm. A practical approach for the computation of Bayes factors from the simulation output is also developed.

• W. S. DESARBO, Y. KIM AND D. K. H. FONG (1999). **A Bayesian Multidimensional Scaling Procedure for the Spatial Analysis of Revealed Choice Data.** *Journal of Econometrics*, 89, 79-108.

The authors present a new Bayesian formulation of a vector multidimensional scaling procedure for the spatial analysis of binary choice data. The Gibbs sampler is gainfully employed to estimate the posterior distribution of the specified scalar products, bilinear model parameters. The computational procedure allows for the explicit estimation of a covariance matrix which can accommodate violations of IIA due to context effects. In addition, posterior standard errors can be estimated which reflect differential degrees of consumer choice uncertainty and/or brand position instability. A marketing application concerning the analysis of consumers' consideration sets for luxury automobiles is provided to illustrate the use of the proposed methodology.

• R. E. MCCULLOCH, N. G. POLSON AND P. E. ROSSI (2000). **A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters.** *Journal of Econometrics*, 99, 173-193.

The authors present a new prior and corresponding algorithm for Bayesian analysis of the multinomial probit model. Their new approach places a prior directly on the identified parameter space. The key is the specification of a prior on the covariance matrix so that the (1,1) element if fixed at 1 and it is possible to draw from the posterior using standard distributions. Analytic results are derived which can be used to aid in assessment of the prior.

CHANGES OF DIRECTIONS

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

Hello!, our names are Maria Eugenia and Javier, and we are in charge of the students' corner this year. We are Ph.D. students at University of Valencia, at the same time that we work as assistant professors at Miguel Hernandez University from Elche (Spain). First of all we want to thank Fabrizio Ruggeri for offering us the possibility to collaborate on the ISBA bulletin, we are sure it will be a very interesting and productive experience for us.

During this year we would like to continue offering information about research developed by students, Ph.D. dissertations, papers, meetings, etc. Moreover, we are keen on communicating how Bayesian statistics is taught in the Ph.D. programs, in different universities from different countries.

As in previous editions, we will present information about personal experiences on applications and opportunities for young statisticians in the field of industry, organizations, laboratories, financial institutions, etc. We also consider interesting the possibility of publishing different experiences in job interviews, and also what kind of statisticians the job market is looking for.

To obtain all these goals, and to get that this section collects as much information as possible, we also need some help and collaboration from all of you. That is the reason why we encourage students to send us their comments about some of the topics above and other suggestions.

In this issue we present a perspective on the current job market. Dr. Maino tells us what type of work could do a statistician in the International Monetary Fund. We conclude the section with the abstract of the thesis of Dr. Streftaris.

Rodolfo Maino[1]
International Monetary Fund, N.W., Washington.
rmaino@imf.org

What does it take to work at the Fund?

For those of you -young and not so- who are seeking jobs in economics, finance and statistics, there is an open door at the International Monetary Fund (IMF, or simply the Fund). The Fund is motivated to widen its recruitment by targeting job seekers with either relevant work experience or preparation in macroeconomics and political economy, finance, statistics, business administration, environmental studies, income distribution, etc. Experienced economists with 5-15 years of experience in economic policy issues, coming from a Ministry of Finance, a Central Bank, or a private financial institution will surely find quite appealing to work at the Fund. Concomitantly, work in academia, research institutions, or other international organizations are also considered a good prerequisite to apply to the Fund. However, to find out, more precisely, what you can offer to the Fund we need first to understand what is the main mission of the Fund and the main avenues (programs) by which you might take a step into it. So let me be more specific.

➤ *What does the Fund do?*

Surveillance is the most important responsibility of the Fund. This fact follows from its original role as the overseer of the Bretton Woods system of fixed-but-adjustable exchange rate parities. Although the system collapsed in 1971, the Fund's focus on members' economic policies still exerts a noticeable influence on the behavior of exchange rates and on the international monetary system. Along these lines, the Fund reviews economic performance of each of its members' countries, prepares an examination of Capital Markets and the World Economic Outlook (the famous WEO for those of you already initiated in economics) and assess members' financial systems and institutions.

➤*What does the Fund offer?*

One of the most interesting -and central- aspects of a Fund economist's responsibilities is the concept of a mission. The latter consists of a small team (3-5 economists) visiting a member country for couple of weeks. An economist usually travels, on average, two to four times a year.

[1]The opinions expressed here are my own and do not necessarily reflect the views of the International Monetary Fund, its staff, its management, or its Executive Board.

After a comprehensive analysis of the latest and relevant data, the team reviews economic developments and policies with the government authorities and, if the country wishes to borrow from the Fund, a financial program is negotiated. The mission then returns to Washington Head Quarters where it prepares a report, to be submitted to and discussed by the Fund's Executive Board. As advertised by the Fund, a keen analytical mind, strong quantitative skills, good teamwork, patience, tact, and stamina are some of the many requirements demanded of staff members during a mission. To undertake these tasks, the Fund welcomes experienced economists as well as young researchers. Among the several paths that lead to a workplace at the Fund, we count the Economist Program, Internship, and a Research Assistant Program.

• *Economist Program*: This is the usual gateway into the Fund for economists entering the organization shortly after graduation. Participants join the Program for a two-year period, after which an appointment to become permanent ("regular") staff might be offered, based on the performance during the previous period. During the program, participants are given two assignments, each lasting one-year. One assignment is in an area department dealing with country-specific issues. The other is likely to be in a functional department, focusing on fiscal, monetary, balance-of-payments, debt and other issues. Participants normally take part in at least two missions during the program. In that endeavor, they work closely with experienced staff members and are given significant responsibility in carrying out operations-related analysis and research. Usually, over 95 percent of these economists join the regular staff, thereby contributing to the work of their team, including participation in missions. They also have access to in-house training seminars (on economics, Fund policies and operations, finance, management, computer and writing, and also language instruction) and external conferences to enhance their professional skills and abilities.

• *Internship*: The Fund offers internships to graduate students each summer to carry out a research project under the supervision of an experienced economist. These projects assigned to summer interns differ from year to year depending on the work program of the intern's division. Interns must set up a paper outlining the results of the research activity assigned to them. Papers of high standard may be published internally. From over 1,000 summer intern candidates worldwide only 35-40 are accepted. Departments typically make the selection based on the closeness of the match between the training and experience of the interns and the demands of the project under consideration. The latter are therefore likely to be close to the interns' own research interests. The research topics are applied-oriented and they are linked to the Fund's main work. They cover a diverse agenda of economic topics, such as: Bank Restructuring, Explaining Exchange, Rate Volatility, Long-Run Fiscal Policy, The EU/EMU Accession for Exchange Policy During the transition Period, The Measurement of International Capital Flows, Imperfect Competition, Asymmetric Information, and Regulation in Financial Markets.

• *Research Assistant Program*: This program is planned for BA students with a superior academic record recently graduated from leading universities who would like to gain useful work experience before pursuing further studies or moving on to other employment opportunities. Appointments are limited to two years and are expected to leave the Fund at the end of the two-year period. Participants mainly support the operational and policy work of Fund economists by being involved in:
• Research, collect and compile information
• Maintain and update quantitative and qualitative economic, financial, or statistical databases
• Process, consolidate and transform data sets within and between databases using statistical and/or econometric techniques
• Evaluate economic, financial or statistical relationships in databases
• Analyze economic time series data
• Design systems to facilitate interfaces or to transfer data between external and in-house databases

➤ *Is there a room for Bayesians?*

You might wonder whether there is a place for those of us already convinced about the advantages of Bayesian theory and methods. Yes of course. The study of human, economics and institutional affairs allows for direct probability statements (such as comparing probability statements of different alternatives to find out which one is more effective). Therefore, applied research and the overview of economic performance make possible the calculation of probabilities, in a Bayesian sense, of future observations. Within that framework it becomes natural to incorporate a priori information from previous experience and experiments into overall conclusions. Although the subjective approach is not yet as popular at the Fund as the frequentist alternative, I believe there is a vast field of application of Bayesian procedures in applied economics and finance. It is just a matter of time.

If, after reading these paragraphs you are committed to apply, please review other qualifications and requirements to be admitted at `www.imf.org`. Remember that candidates should initially submit their CV or a Fund Application Form, indicating education and work experience. The Fund may later also request samples of written work. Candidates are screened in light of current and potential vacancies. Fund recruiters visit universities, public sector organizations, and major economics conferences in many countries throughout the year to contact potential applicants. If an individual's qualifications match the requirements of a current or expected vacant position, preliminary interviews are arranged, often in the candidate's own country or the country where he or she is currently working. If these go fine, the promising applicant is likely to be invited to Washington for final interviews with a panel of senior Fund economists in the interested department(s). Good Luck!!

Dr. George Streftaris
Actuarial Mathematics and Statistics, Heriot-Watt University
`georges@ma.hw.ac.uk`

Bayesian methods for Poisson models

Advisors: Dr. Bruce Worton and Dr. Tom Leonard

To account for overdispersion in count data, that is variation in excess of that justified from the assumed model, one may consider an additional source of variation, by assuming that each observation, Y_i, $i = 1, \ldots, m$, arises from a conditionally independent Poisson distribution, given its respective mean θ_i, $i = 1, \ldots, m$.

We review various frequentist methods for the estimation of the Poisson parameters θ_i, $i = 1, \ldots, m$, which are based on the inadmissibility of the usual unbiased maximum likelihood estimator, in terms of the associated risk in dimensions greater than two. The so-called shrinkage estimators adjust the maximum likelihood estimates towards a fixed or data-determined point, abandoning unbiasedness in favour of lower risk.

Inferences for the parameters of interest can also be drawn employing Bayesian methods. Conjugate models are often adopted to facilitate the computational procedure. In this thesis we assume a non-conjugate log-normal prior distribution, which allows for more dispersion in the Poisson means and can also accommodate a correlation structure. We derive two empirical Bayes estimators, which approximate the posterior mean. The first is based on a linear shrinkage rule, while the second employs a non-iterative importance sampling technique. The frequency properties of the two estimators in terms of average risk are assessed and compared to other estimating approaches proposed in the literature.

A full hierarchical Bayes analysis is also considered, assuming both informative and vague prior distributions at the lower stage of the hierarchy. Some analytical posterior inferences, based on approximations are obtained. We then employ stochastic simulation techniques, suggesting two Markov chain Monte Carlo methods which involve the Gibbs sampler and a hybrid strategy. They rely on a log-normal/gamma mixture approximation to the full conditional posterior distribution of the parameters θ_i, $i = 1, \ldots, m$. The shrinkage behaviour of the hierarchical Bayes estimator is explored, and its average risk is examined through frequency simulations. Examples and applications of the considered methods are given throughout the thesis.

NEWS FROM THE WORLD

by Antonio Lijoi
lijoi@unipv.it

✻ *denotes an ISBA activity*

➤ Events

Workshop on Recent Advances in Bayesian Econometrics. *June 14-15, 2001, Marseilles, France.*

The aim of this workshop is to explore the last advances of Bayesian Econometrics. It builds on the tradition initiated by the two previous conferences organized by the members of the LEA (European Associated Laboratories, CORE Louvain and GREQAM Marseilles). These two conferences were organized in Marseilles in 1992 and in 1996 over the Rhine (River Boat Conference). They resulted in the edition of two special issues of the *Journal of Econometrics*. This year, the organisers propose to discuss the contributions of Bayesian methods preferably in three fields : model choice, Monte Carlo methods and panel data. The workshop's website is durandal.cnrs-mrs.fr/bayes/.

Workshop on Bayesian Models in Medicine. *July 1, 2001, Cascais, Portugal.*

Bayesian networks (also known as causal probabilistic networks) with their associated methods have now been around in biomedicine for more than a decade. They have become increasingly popular for representing and handling uncertain knowledge in medicine. Almost simultaneously, the use of Bayesian statistics has increased in popularity in medicine. This workshop aims to bring together researchers in these fields in order to assess the current state of the art, to identify obstacles for progress and to determine future research directions. Extended versions of the best papers of the workshop will be published in a special issue of the journal Artificial Intelligence in Medicine. The deadline for the submission of papers is May 1, 2001. Web page of the workshop: www.csd.abdn.ac.uk/~plucas/aime01-ws.html

Summer School on Limit Theorems in Statistics. *July 2-21, 2001, Torgnon (AO), Italy.*

The purpose of the course is to introduce graduate students to some relevant topics of asymptotic theory in statistics. The course will consist of a series of lectures held by Prof. Y. Rinott and of tutorials held by Dr G. Zitkovic. Application forms should be sent to Prof. D.M. Cifarelli, Università Bocconi – Istituto di Metodi Quantitativi, V.le Isonzo 25, 20135 Milano. Deadline for submission of applications is May 21, 2001. Full detailed program of the course and application form to be submitted may be found at www.uni-bocconi.it:1080/imq/corsoestivo2001/

Taipei International Quantitative Finance Conference. *July 3-5, 2001, Taipei, Taiwan.*

The conference is organized by the Institute of Statistical Science, Academia Sinica, Taiwan. It will focus on recent developments in the field of Quantitative Finance, including economic and statistical modeling, stochastic financial theory, model estimation, financial data analysis and statistics, derivatives, market risk and numerical simulation. The deadlines for the submission of papers and for pre-registration are April 15, 2001 and May 31, 2001, respectively. Web page of the conference: www.stat.sinica.edu.tw/tiqfc2001/english/main.htm

Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering. *August 4-9, 2001, Baltimore, MD, USA.*

The annual Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt) promotes the development of theoretical and applied aspects of inductive logic. This workshop is dedicated to Dr. Richard Threlkeld Cox. The six-day agenda includes a one-day tutorial session, invited lectures and papers, contributed papers and poster presentations. Contributed papers are requested on topics including applied Bayesian inferencing and maximum entropy methods, inductive logic theory, and theoretical physics. Deadline for receipt of abstracts is 1st April 2001. Preliminary workshop information for MaxEnt2001 is available at www.jhuapl.edu/maxent2001/.

Summer School on Stochastics and Finance. *September 3-7, 2001, Barcelona, Spain.*

The purpose of this advanced course is to provide a forum to people interested in the recent developments of Mathematical Finance. The main focus will be on hedging and, as a second topic, alternative models. This advanced course will have a series of 5 one hour and a half lectures on "Recent Developments in Hedging" held by Prof. I. Karatzas. There will also be two other smaller short courses on "Lévy systems in Finance" by Prof. D. Madan, and on "Lévy based dynamic models for financial economics" by Prof. O.E. Barndorff-Nielsen. The level of the courses will be at the postgraduate level and some basic concepts of mathematical finance will be assumed as well as a good level of stochastic analysis. We expect to offer some partial grants for young participants from EC countries. Further details may be found at he web page: `www.imub.ub.es/events/sssf/index.html`

First Annual Conference on Business and Industrial Statistics. *September 17-18, 2001, Oslo, Norway.*

The European Network for Business and Industrial Statistics (ENBIS) was founded in 2000 as a network of people from industry and academia from all European nations interested in applying, promoting and facilitating the use of statistics in business and industry.

The purpose of this first annual conference is to create a forum for users of Statistics to get together, share ideas and network. Papers will be presented showcasing a broad spectrum of applications and generate discussion about the use of statistics in a wide range of European business and industrial areas. The deadline for contributing papers is May 20, 2001. Further information at the web page: `www.ibisuva.nl/ENBIS/`

Summer School on Advanced Computational Methods for Statistical Inference. *September 17-21, 2001, Luminy, France.*

The purpose of this course is to bring together senior researchers on the topic of recent advances in computational methods, with applications to complex statistical problems, like those encountered in econometrics, signal processing and finance, and to introduce Ph.D. students and junior researchers to these methods. Courses will be held in the mornings, advanced talks in the afternoon and round-tables in the evenings. The summer school is funded by the European Union's TMR network "Statistical and Computational Methods for the Analysis of Spatial Data". The deadline for registration is July 1, 2001. Further information at the web page: `tsi.enst.fr/~moulines/SummerSchool/presentation.html`

Euroworkshop on Statistical Modelling - Nonparametric Models. *November 1-4, 2001, Munich, Germany.*

The Euroworkshop on Statistical Modelling is a project which is sponsored by the European Commission (CORDIS) in the programme High Level Scientific Conferences. The workshop on Nonparametric Models is the second event in a series of three workshops with the topics Mixed Models, Nonparametric Models and Model Diagnostics. Coordinator of the entire project is Göran Kauermann. It is planned that 1/3 of the participants are young researchers. Young Researcher can be supported by the European Commission in form of travel grants and free accommodation. Web page of the Euroworkshop: `www.stat.uni-muenchen.de/euroworkshop/2001.html`

The Eighth Latin American Conference on Probability and Statistics. *November 12-16, 2001, La Habana, Cuba.*

For more information about the conference, to be held at the University of La Habana, please visit the website `www.uh.cu/eventos/clapem/ehome.htm` or contact Gonzalo Perera, the Chair of the Program Committee at `gperera@fing.edu.uy`, or Pablo Olivares, Chair of the Local Organizing Committee at `clapem@matcom.uh.cu`.

➤ Research Opportunities

Postdoctoral fellowships at EURANDOM.

The European research institute EURANDOM carries out research in mathematical statistics, probability theory,

operations research and their applications. At the moment EURANDOM has vacancies for *Postdoctoral Fellows* in the area of mathematical statistics and its applications in industry. The post-docs will work both on an industrial project and on theoretical projects. One of the industrial projects is a 4-year project with a leading manufacturer of copying machines. The goal of this project is to develop advanced methodologies for on-line failure predictions of copying machines at remote places. Candidates for a postdoctoral fellowship are encouraged to send a letter of application, together with a curriculum vitae with full educational details, three letters of recommendation, a list of publications, pre-prints of selected papers or a thesis abstract to: Prof.dr. W.Th.F. den Hollander, Scientific Director EURANDOM, P.O. Box 513, 5600 MB Eindhoven, the Netherlands Phone + 31 40 247 8100; fax + 31 40 247 8190; e-mail: `office@eurandom.tue.nl`.

Web page for information about EURANDOM and the research programmes: `www.eurandom.tue.nl`

Research student funding opportunities at the University of Reading.

At the University of Reading, Department of Applied Statistics, there are four PhD studentships available for research in the field of statistical genetics and genomics.

– BBSRC quota studentship for any one of a large number of projects in *statistical genetics*. Supervisors: Prof. David Balding, or Dr Mike Denham, or Dr John Whittaker.

– BBSRC studentship for the project: *Statistical Approaches for the Modelling of Gene Expression Using Array Data*. Supervisors: Dr Mike Denham and Dr Simon Andrews (School of Animal and Microbial Sciences).

– BBSRC studentship for the project: *Investigating the Genealogies of Small Populations*. Supervisors: Prof. David Balding and Dr Mark Beaumont (Animal & Microbial Siences).

– EPSRC studentship for the project: *Extracting and Representing Information from Array Images of gene and Protein Expression*. Supervisor: Dr Mike Denham.

Further details are available at `www.rdg.ac.uk/statistics/phd/`

Research Assistant position at the University of Strathclyde and the University of Glasgow

The successful candidate for this appointment, which is for a period of up to 36 months, will be employed jointly by the University of Strathclyde and the University of Glasgow for the project"*Spatial Stochastic Modelling and Optimisation of Broadband Telecommunications Networks*". This project is funded by EPSRC and will use methods of probability theory, stochastic geometry, spatial processes and modern simulation techniques for studying, modelling and optimising of complex telecommunications systems. Applications (Ref: R12/01) with 3 copies of a full CV giving a full list of scientific publications and the names and addresses (including email addresses where appropriate) of two academic referees should be sent to Dr Sergei Zuyev, Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH. Applications closing date: May 1, 2001. `E-mail: sergei@stams.strath.ac.uk`

➤ Awards and Prizes

2001 Spiegelman Award Nominations.

The American Public Health Association invites nominations for the 2001 Mortimer Spiegelman Award honoring a statistician aged 40 or younger who has made important contributions to public health. The award recognizes outstanding contributions to the field of health statistics. Please send a nominating letter and the candidate's CV to: Peter Bacchetti, Chair, Spiegelman Award Committee, Department of Epidemiology & Biostatistics, Box 0560 University of California San Francisco, CA 94143-0560. Nominations must be postmarked no later than May 15, 2001. Web page: `chanane.ucsf.edu/biostat/spiegel.html`

INTERNATIONAL SOCIETY FOR BAYESIAN ANALYSIS

Web page: *http://www.bayesian.org/*

EDITORIAL BOARD

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: *http://www.iami.mi.cnr.it/isba*