THE ISBA BULLETIN

Vol. 18 No. 2 June 2011

The official bulletin of the International Society for Bayesian Analysis

A MESSAGE FROM THE PRESIDENT

THE ERA OF BIG DATA

- Michael I. Jordan -
ISBA President, 2011
jordan@stat.berkeley.edu

The era of "Big Data" is upon us. It is a rare week in which one does not see an article in a major international magazine or newspaper on the phenomenon of Big Data. One sees discussions of terabytes, petabytes, exabytes and zettabytes. (Megabytes and gigabytes seem to be viewed as old-fashioned.)

Beyond discussions of mere size (and rates of growth in size), the focus is increasingly on inferential issues. Past research on databases and information retrieval is viewed as having focused on mere storage, lookup and search, whereas the opportunity now is to discover new phenomena and to "learn." There is also an increasing awareness that such opportunities are counterbalanced by dangers of "discovering" phenomena that are not real.

Big Data issues arise both in science and technology. In science, massive streams of data have become the norm in areas such as astronomy, high-energy physics, ecology, genetics and molecular biology. In technology, there is a new drive towards personalization, based on the increasing availability of data on fine-grained aspects of human behavior; the analysis of such data is viewed as permitting the development of new services that are tailored to individuals.

There seems to be a growing awareness that the challenges involved in meeting Big Data problems are interdisciplinary. Data analysis in the Big Data regime requires joint consideration of systems issues (how to store, index and transport data at massive scales; how to exploit parallel and distributed platforms), statistical issues (how to cope with errors and biases of all kinds; how to develop models and procedures that work when both n and p are astronomical), and algorithmic issues (how to perform computations using resources that scale as linear or sub-linear functions of n and/or p). And of course there are legal, commercial and social issues.

One sees responses to these challenges at all levels of society. Companies are increasingly looking to hire people who have expertise in data analysis. . . . *Continued on page 2.*

In this issue

MESSAGE FROM THE PRESIDENT, *Continued from page 1.* . . . While the job descriptions sometimes refer to "statistics," they often refer to "data analytics," "data science," "data mining" and "machine learning." In any case, there seems to be an increasing awareness that many of the fundamental underlying problems are statistical in nature. Funding agencies are paying attention to Big Data, and there is steady growth in the numbers of workshops and conferences devoted to Big Data issues.

What is the role of Bayes in all of this? To date, I would venture to say "unfortunately rather little." When discussions eventually turn to methodology—in the job descriptions or the conference talks—one sees discussion of ideas such as hierarchical clustering, K-means, the support vector machine, boosting, random projections, nearest neighbor and the Lasso. Rarely is Bayes mentioned.

I am occasionally invited to give talks in workshops on Big Data, and my somewhat contrarian nature leads me to give talks on Bayesian approaches. This is met with immediate skepticism—are you telling us that we should be rolling out MCMC methods in the Big Data world, when such methods are often too slow for old-fashioned small-data problems?

I am skeptical about this skepticism, but I am rarely able to formulate a very convincing response. Unfortunately, the Bayesian literature has not yet taken Big Data very seriously, and we simply do not know what kinds of Bayesian methodology might be viable at massive scales.

Another criticism goes as follows: Bayes is principally concerned with error bars rather than point estimates, and in the realm of Big Data the error bars are negligible, so why do we need Bayes? This argument is easier to rebut. In analysis of Big Data one is rarely concerned with quantities such as population means; rather one is concerned with small sub-populations, and concerned with the tails of distributions. Moreover, Big Data analysis generally involves many error-prone steps, and the propagation and control of uncertainty is critical.

In general I believe that Bayes has several natural advantages in the world of Big Data. Let me note four such advantages: (1) Analyses of Big Data often have an exploratory flavor rather than a confirmatory flavor. Some of the concerns over family-wise error rates that bedevil classical approaches to exploratory data analysis are mitigated in the Bayesian framework. (2) In the sciences, Big Data problems often arise in the context of "standard models," which are often already formulated in probabilistic terms. That is, significant prior knowledge is often present and directly amenable to Bayesian inference. (3) Consider a company wishing to offer personalized services to tens of millions of users. Large amounts of data will have been collected for some users, but for most users there will be little or no data. Such situations cry out for Bayesian hierarchical modeling. (4) The growing field of Bayesian nonparametrics provides tools for dealing with situations in which phenomena continue to emerge as data are collected. For example, Bayesian nonparametrics not only provides probability models that yield power-law distributions, but it provides inferential machinery that incorporate these distributions.

But there remains the elephant in the room: will the traditional tools of Bayesian posterior inference be able to cope with arbitrarily large data streams? To this I would respond that we simply do not know. Although one can certainly imagine that algorithms such as MCMC and sequential Monte Carlo will scale increasingly poorly as n and p increase without bound, it is also conceivable that this will not happen in general. Perhaps the sharing of statistical strength implemented by Bayesian hierarchical models will lead to implicit computational efficiencies, at least for some aspects of the posterior distribution (ideally, for the inferences that matter). There are research questions here that have not begun to be touched.

The recent growth of activity in parallel and distributed implementations of posterior inference algorithms (e.g., GPUs) is encouraging in this regard. But we should not view parallelism as a silver bullet; it will not suffice to scale Bayesian inference to massive data sets. We should instead hope that this line of research will lead to increased understanding of how posterior inference algorithms behave at larger scales, hopefully pointing the way towards further modeling and algorithmic developments.

Finally, I wish to make a somewhat philosophical note, one which is as applicable to non-Bayesian statistics as to Bayesian statistics. Suppose that I am a statistician faced with an inference problem involving a possibly large data stream. Sallying forth with a fast computer and the R package under my arm, I am pleased to be given more and more data, up to a certain point, beyond which I start to become concerned. I eventually throw up my hands, unable to

do much of anything. This seems odd in a certain philosophical sense. If data are our principal resource, why should having too much data cause such embarrassment? Shouldn't I be able to "throw away" or "reduce" data as they accrue and still maintain a certain quality of statistical inference within a given computational budget? Indeed, despite my limited computational budget, shouldn't I achieve a growing quality of statistical inference as data accrue? (Note that the act of data reduction involves a computational cost, which must be taken into account in the budget.) Probably the answer is yes, but we are very far in the current state of our discipline from being able to provide such guarantees.

In short, I would propose that we begin to think about how to do Bayesian data analysis in the context of data sets of *arbitrary* size; let's aim to overcome our embarrassment at the riches of massive data once and for all.

The Big Data problem is here to stay, and Bayesians should begin to think Big. ▲

A Message from the Editor

Manuel Mendoza
mendoza@itam.mx

In this Bulletin our President, Michael Jordan, calls our attention on the increasing number of problems where the data, far from being scarse, is massive. The need for statistical procedures, capable to deal with huge amounts of complex data to produce useful, effective and timely results is discussed in detail.

Also in this issue, we have a report by Paul Demian on the Conference that was recently held in Creete, in honour to Prof. Adrian F.M. Smith. More than a hundred colleagues, many of them coauthors and/or former students of Sir Adrian Smith, gathered in the Aegean Sea in an impressive academic tribute to one of our leaders.

Brief notes from the sections of ISBA describe their recent activities and call for an active participation to those who may be interested in these areas of research.

As every quarter, you will also find the usual sections of the Bulletin. In particular, I call your attention to the Annotated Bibliography Section where the Editor, Beatrix Jones, asked Prof. Anthony Pettitt to write an article on the main contributions of the late Julian Besag. The article includes an impressive list of areas where Prof. Besag contributed decisively. Our Software Editor, Alex Lewin, invited Jong Hee Park to introduce some recent developments of the R package MCMCpack. In addition, our Student's Corner Editor, Luke Bornn, continues his quest and poses one more question to his panel of distiguished colleagues. Again, the result is interesting and revealing.

Finally, I want to encourage all members of ISBA to contribute to the Bulletin with their suggestions, manuscripts and announcements. Please do not hesitate to contact me or any member of the Editorial Board. ▲

BAYESIAN ANALYSIS - A MESSAGE FROM THE EDITOR

UPDATE FROM BA

Herbie Lee
Editor-in-Chief
herbie@ams.ucsc.edu

The big event at Bayesian Analysis this quarter is that we are saying goodbye to Angelika van der Linde, who is stepping down after many years as our Production Editor. Angelika has been crucial in helping us produce quality issues and we need to thank her for her efforts and contributions. We also are welcoming in Kary Myers, who is our new Production Editor, and we are excited to have her join our team.

The June issue of BA features a paper on covariance structures for multidimensional arrays by Peter Hoff. Use of the Tucker product provides a practical framework for approaching the problem. Additional discussion and perspective appears in discussions by Genevera Allen and Hedibert Lopes. This issue also contains other fine articles in computational statistics, bioinformatics, stochastic processes, and nonparametrics.▲

FROM THE PROGRAM COUNCIL

ISBA NEWS

Igor Pruenster
igor@econ.unito.it

Call for special topic sessions at ISBA 2012

The Program Committee of ISBA 2012 invites submissions of proposals for special topic sessions. Each session consists of 4 talks of 25 min with a common theme. The proposal should include the title of the session, a brief description (session abstract), and the list of 4 speakers and should be sent to isba2012@e.u-tokyo.ac.jp by September 18th, 2011.

The ISBA 2012 World Meeting – the premier conference of the International Society for Bayesian Analysis (ISBA) – will be held in Kyoto, Japan, from June 25 to June 29, 2012. Preliminary program can be found at http://www2.e.u-tokyo.ac.jp/~isba2012/.

Details about support for junior presenters will be finalized in late 2011. Please check the conference website for updates.▲

ISBA - SECTIONS

OBJECTIVE BAYESIAN SECTION

Jim Berger
Chair
berger@stat.duke.edu

OBayes 2011. This workshop, formally the 2011 International Workshop on Objective Bayes Methodology, was held in June 11-15, 2011 at East China Normal University in Shanghai, China. The conference had roughly 100 participants, half from China and half from other countries. It was designated as one of the principle events in ECNU's 60th anniversary celebration, appropriate as the university started the first mathematical statistics department in China. ECNU also used the occasion to inaugurate a new Academy of Applied Statistical Sciences, an academy in which Dongchu Sun (Program Chair of the OBayes section) will play a central role.

Since students in China typically receive only modest exposure to Bayesian analysis, the conference began with a full day of tutorials. There were then a total of 21 invited talks with discussants, followed by lively floor discussion; many of the talks are posted at http://www.sfs.ecnu.

edu.cn/Obayes2011/index.html. Other conference highlights included a well-attended poster session with the requisite refreshments (best student poster winners were Xiaojing Wang from Duke U., Anchu Xu from ECNU, and Chang Xu, U. Missouri); wonderful meals, including an awesome banquet with hilarious after dinner remarks by Ed George; and a spectacular tour of Shanghai - including a delightful boat cruise that had no problems whatsoever; perhaps the "Bayesian conference boat excursion curse" has been lifted!

The program committee and local organizers of the workshop - led by Dongchu Sun and Yincai Tang - did a wonderful job, and ISBA played a valuable role in handling the registration and promoting the workshop.

This was the first formal meeting of the Objective Bayesian Section of ISBA and, I believe, was the first Bayesian meeting ever held in China; we hope there are many more to come!

OBayes 2013. Please send me any suggestions you might have concerning the location and timing of the 2013 OBayes meeting.▲

NONPARAMETRIC BAYES SECTION

Ramses Mena
Program Chair
ramses@sigma.iimas.unam.mx

Precisely one year ago, in June 2010, Stephen Walker wrote a brief article for the Bulletin announcing the Nonparametric Bayes Section of ISBA. Since then, the nonparametric community has been working hard and while this note is been written, the **8th Workshop on Bayesian Nonparametrics** is taking place in Veracruz, Mexico.

The objective of the workshop is to bring together experts and young talented scientists devoted to the study and application of Bayesian nonparametric techniques. It was organized under the auspices of the Mexican Statistical Association (AME) and the International Society for Bayesian Analysis (ISBA).

Here, we had the opportunity to hear 19 invited presentations as well as 36 contributed talks together with 4 tutorials and to enjoy a superb view of the portuary activities from the wonderful lounge located at the top floor of the hotel where the posters session was organized. Around 150 well stablished researchers, young scientists and students from 15 different countries and 67 institutions attended the workshop in a warm and friendly environment.

Finally, I would like to thank all the people, institutions and societies who support us while organizing the workshop and as for the section, please send me any comments, suggestions or questions you might have regarding the activities of the section. It will be great to hear from you.▲

ADRIAN F.M. SMITH Conference

HIERARCHICAL MODELS AND MARKOV CHAIN MONTE CARLO

JUNE 2 TO 5, 2011. CREETE, GREECE.

Paul Damien
Paul.Damien@mccombs.utexas.edu

Set against the back drop of sunny skies, cool breezes, and the bluish green waters of the Cretan Sea, the Knossos Royal Village was the venue for the conference honoring Adrian Smith, a Bayesian statistician whose many path-breaking contributions to the theory and practice of statistics continue to impact academic research and industrial practice world-wide.

Conference Facts

Approximately 125 to 130 people attended the event, which excluded parallel sessions, thus forcing Adrian to participate in all the talks! There were a total of 32 talks, each 30 minutes long, and 17 poster presentations. The speakers were encouraged to address broad statistical ideas, focusing, if possible, on the twin themes of Hierarchical Models and Markov Chain Monte Carlo -the banner heading for the conference. To ensure a leisurely pace, the sessions were interspersed with multiple coffee breaks, culminating in an early afternoon siesta at 15:30 hours, followed by dinner at 20:00 hours. On the final day, after the last lecture, several group photos were taken by Manuel Mendoza-Ramirez and Eduardo Gutierrez-Pena, prompting Adrian to quip, "how many Mexicans does it take…". The banquet dinner was held by the main swimming pool. The highlights of the after-dinner show included a humorous Monty Python skit by Linda Sharples, Jon Forster, Nick Polson, and Peter Mueller. The inimitable Ed George proceeded next with a series of (by his own admission) ´´tasteless jokes", and concluded by poignantly noting that "we all miss Adrian". A brief slide show of Adrian's photographs, set to the tunes of the Irish tenor John McCormack, was then followed by a roasting of Adrian by that natural-born (Bayesian) stand-up comedian John Deely. Finally, the moment that we had all been waiting for: Adrian took the podium. His flawless and inspiring 5-minute delivery confirmed, once again, why he holds a singular position in the advancement of statistical science. The entire show was filmed by Wes Johnson. The film will be posted on the website www.afmsmith.com in July.

The Talks

Applying a valuable Adrian Smith dictum, namely the importance of brevity in communication, the welcoming remark was made by Petros Dellaportas, who kicked off the event by stating, "Welcome". He then introduced the first speaker David Spiegelhalter. David's aptly titled talk, ´´Following the leader of the revolution", was wildly entertaining and scientifically informative; it set the tone for the rest of the meeting. The

final speaker was Gareth Roberts. He bookended David's talk by punctuating his own with a series of laugh-out-loud anecdotes and jokes, while cleverly interjecting them with deep insights on convergence issues in MCMC within a hierarchical modeling framework. The speakers list may be accessed at www.afmsmith.com.

The Organizing Committee

Paul Damien, Petros Dellaportas, Jon Forster, Nick Polson, Gareth Roberts, David Stephens, Jon Wakefield▲

ANNOTATED BIBLIOGRAPHY

The Contributions of Julian Besag

James McKeone & Tony Pettitt
jp.mckeone@student.qut.edu.au
ja.pettitt@qut.edu.au

This article is joint work with James McKeone. James is just starting a PhD at QUT which involves a problem from neurology and the modelling of action potentials recorded at different stimulus levels on an 8 by 16 regular array of 128 electrodes. There are some interesting spatial data with considerable structure. A Bayesian approach is planned, and James enjoyed very much reading several of Julian's papers.

Tony Pettitt

News of Julian's death was very sad but gave me the opportunity to reflect on his immense contribution to modern statistics and my personal contact with him.

I must have first met Julian at one of the monthly RSS read paper meetings which I attended regularly, getting the train down to London from Nottingham or Loughborough (where Julian's father had been a lecturer). Later, I came face to face with him over an interview panel in 1982 for a lecturing job at Durham. Julian asked me why I had not had a paper read to the RSS and, somewhat tritely, I replied that I had been in Australia for two years. I do not think the answer was well received.

In 1991 I had a position in Australia at QUT. I organized a statistical computing and biometrics meeting at Coolangatta on the beach on the Queensland / New South Wales border. I had invited Sandy Weisberg and Julian as the statistical computing invited international speakers. Julian gave a talk on Gibbs sampling and spatial statistics and I remember his carefully drawn overhead transparencies of spatial DAGs. That must have been the first introduction in a public lecture of the those ideas to an Australian audience. It sowed some Bayesian seeds not least with my colleague at QUT, Kerrie Mengersen.

In the remarks below the two read papers of 1974 and 1986 are given prominence. "Spatial interaction and the statistical analysis of lattice systems"(1974) was considered seminal at the time

it was published and more than lived up to this promise. I was present at the 1986 meeting where "Statistical analysis of Dirty pictures"was read. At the end of Julian's presentation he received a great acclamation, even though we never saw the "dirty pictures" or perhaps relieved that we had not. He gave the talk with the confidence of someone who knew absolutely that he had a winner (he was the most highly cited mathematical scientist of the 1980s). And that was always the way Julian gave a seminar, with immaculate preparation and confident delivery.

Conditional Framework

- J. E. Besag (1972) "Nearest-neighbour systems and the auto-logistic model for binary data." Journal of the Royal Statistical Society Series B (Methodological) 34:75–83. Considers a conditional framework for spatially dependent random variables on a regular lattice. Finding that in the case of binary homogeneous random variables, the only theoretically plausible model is the now familiar auto-logistic process. Some spatio-temporal models are investigated but they are noted to be of limited practical appeal.

- J. E. Besag (1974) "Spatial interaction and the statistical analysis of lattice systems."Journal of the Royal Statistical Society Series B (Methodological) 36:192–236. Further develops the ideas seeded in Besag (1972). The focus of the paper is on binary and Gaussian models with applications in plant ecology. A simple, alternate proof of the Hammersley-Clifford theorem is presented assuming the *positivity* assumption. The theorem may be used to specify specific spatial schemes under modest assumptions. Methods for inference are also presented along with a coding technique for model fitting that is easy to implement and has a sound theoretical basis, giving a true but inefficient likelihood. The paper continues to argue the merits of modelling under the conditional framework as opposed to the alternative joint framework.

  The paper was a read paper at the RSS and these are some of the discussants' published comments.

  Professor D. R. Cox (Imperial College): The paper is original, lucid and comprehensive. The topic is important and notoriously difficult. It is a pleasure to congratulate Mr Besag.

  Professor P. Whittle (University of Cambridge): Mr Besag has given us a substantial, interesting and most useful paper. It collects together a great amount of material and clarifies a number of points admirably. For example, the Hammersley-Clifford theorem now appears very much more accessible.

  Dr J. M. Hammersley (University of Oxford): I am much impressed by Mr Besag's very interesting and far-ranging paper on spatial interaction. The variety and flexibility of his treatment will greatly contribute to our understanding of these very difficult problems of co-operative phenomena.

  Professor M. S. Bartlett(University of Oxford): I regret that my absence in Australia has prevented my coming to this meeting, but it is indeed a pleasure to send my congratulations to the author for this very comprehensive and valuable paper on nearest-neighbour lattice models, containing both exposition of basic theory and of statistical techniques of analysing spatial interaction. I find his discussion of auto-normal and auto-logistic model-fitting especially interesting and useful.

  These comments indicate the substantial standing and seminal nature of the paper. The discussion points to the growing awareness of graphical modelling as a tool in statistics.

- J. E. Besag (1977). "Errors-in-variables estimation for Gaussian lattice schemes." Journal of the Royal Statistical Society. Series B (Methodological), 39:73–78. Some notes on estimation when the variates of primary interest are not observable on a simple Gaussian lattice. Notes on estimation procedure for error-in-variables approximation.

Extensions including Pseudolikelihood

- J. E. Besag (1975) "Statistical analysis of non-lattice data." The Statistician 24:179–195. Introduces the pseudolikelihood approach to estimation of spatial interaction of random variables for irregularly dis-

tributed data points. The method permits more computationally tractable parameter estimation by computing an approximation to the likelihood for a set of random variables.

- J. E. Besag and P. A. P. Moran (1975). "On the estimation and testing of spatial interaction in Gaussian lattice processes." Biometrika 62:555-562. Develops a technique for explicit maximum likelihood estimation for first-order, regular lattice autonormal models. The estimation is computed on an infinite lattice scheme, and on a torus lattice with first-order, isotropic square dependence structure. The coding technique is found to be efficient only for relatively small first order correlation.

- J. E. Besag (1977). "Efficiency of pseudolikelihood estimation for simple Gaussian fields." Biometrika 64:616–618. Adjunct to Besag (1975) and Besag and Moran (1975). Notes on efficiency of pseudolikelihood estimation contrasted to the maximum likelihood approach of the former papers. Besag notes that the psuedolikelihood technique produces "encouraging" efficiency results for the first order isotropic regular lattice schemes.

Monte Carlo based inference

- J. E. Besag and P. Clifford (1991). "Sequential Monte Carlo p-values." Biometrika 78:301–304. The computational burden of the using Monte Carlo methods to assess statistical significance is notably high. This paper presents a method of sequential sampling where samples are drawn until a predetermined number of deviates exceed a benchmark value. The ratio of the number of deviates that exceed the benchmark to the total number simulated is an estimation to the true p-value that is exact when H_0 is true and provides a good approximation when H_0 is false.

- J. E. Besag and P. J. Diggle (1977). "Simple Monte Carlo tests for spatial pattern." Journal of the Royal Statistical Society, Series C (Applied Statistics), 26:327–333. A review of Monte Carlo tests for spatial pattern with an emphasis on preliminary investigation. Concludes that there is no specific test that performs well in every circumstance - rather a general approach is preferred, that even when analytical (asymptotic) distributional theory is known, an MC test is often also worthwhile, and that the nature of MC tests is fairly intuitive and easy to communicate to non-statisticians.

Bayesian spatial modellling and MCMC

- J. E. Besag (1986). "On the statistical analysis of dirty pictures." Journal of the Royal Statistical Society. Series B (Methodological), 48(3):259–302. Assumes the local characteristics of the true image are modelled by a non-degenerate Markov random field. Uses *a priori* information in a Bayesian framework to determine the maximum *a posteriori* (MAP) estimate giving the colouring that possesses overall maximum probability given the data. Introduces the iterated conditional modes (ICM) algorithm to aid in parameter estimation due to the computational difficulty of the new model. The ICM algorithm uses simulated annealing and a Gibbs like sweep though pixels finding the posterior most likely value for the pixel. Examples include some noisy images involving true scenes with a number of colours and show substantial improvements over joint pixel maximum likelihood classifiers.

  For computationally efficient estimation of parameters involved in the Markov random field, the paper recommends maximum pseudo-likelihood.

  This paper was a discussion paper at a RSS meeting and the following comments are of historical interest. The first was most prophetic in terms of impact of the paper and the second shows the seeds of BUGS were germinating.

  Professor D. M. Titterington (University of Glasgow): A crucial feature of tonight's paper is the proposal of a joint probability function . . . To put it bluntly, it could be said that the stylish modelling turns out to be a means to the end of proposing, in the form of the ICM method, a practically feasible, if rather ad hoc algorithm . . . and Professor Besag is to be applauded

for leading UK statisticians into this important area. Maybe his technique's title can be extended eponymously to that of Iterated Conditional Besag Modes to reflect, by its initials, his international and explosive impact on the field! (a contemporary reference to the state of the cold war).

Dr D. J. Spiegelhalter (MRC Biostatistics Unit, Cambridge): A number of aspects in tonight's stimulating paper have parallels in current work on expert systems. Thus in place of the pixels in image analysis, we substitute possible clinical findings and diseases, while the regular lattice structure of the relationships expressed in the random field is replaced by a complex network comprising the local relationships archetypically expressed as subjective 'rules'. . . .

- J. E. Besag, J. York, and A. Mollié (1991). "Bayesian image restoration, with two applications in spatial statistics." Annals of the Institute of Statistical Mathematics, 43:1–59. Interprets spatial data analysis problems as problems in image analysis. Introduction of the conditional autoregression (CAR) model and a rank deficient intrinsic alternate (ICAR or IAR). The methodology emphasizes a Bayesian approach, using a Gibbs sampler to compute the maximum of the posterior marginal probabilities. Much cited in the disease mapping literature.

- J. E. Besag and P. J. Green (1993). "Spatial statistics and Bayesian computation." Journal of the Royal Statistical Society. Series B (Methodological), 55:25–37. Presents development of and a review of MCMC techniques in Bayesian inference. The techniques that are discussed are variance reduction methods, antithetic variable methods, convergence and efficiency, use of auxiliary variables, and MCMC under multimodality. With some application to field experiments.

- J. E. Besag, P. Green, D. Higdon, and K. Mengersen (1995). "Bayesian computation and stochastic systems." Statistical Science, 10:3–66. Basic methodology of MCMC with an emphasis on the Bayesian paradigm and modelling spatially dependent systems. Applications in spatial or temporal aspects of the problems, agricultural field experiments, image analysis.

Algebra of interacting Systems

- J. E. Besag (1981). "On a system of two-dimensional recurrence equations." Journal of the Royal Statistical Society. Series B (Methodological), 43:302–309. Considers two-dimensional recurrence equations of a first order auto-normal scheme on a first-order, isotropic, doubly infinite rectangular lattice. It is shown that the covariance function of a Gaussian Markov random field can be approximately represented in terms of a modified Bessel function that decreases monotonically with distance.

- J. E. Besag and C. Kooperberg (1995). "On conditional and intrinsic autoregression." Biometrika, 82(4):733–746. A summary of Gaussian conditional autoregressions - intrinsic or otherwise. Illustrates how Dempster's algorithm may be used to alleviate the problem of an undesirable marginal variance structure in CAR models. Shows that under the intrinsic CAR model, the parameters may be estimated such that X_i and X_j have constant variance when i and j are neighbours.

- F. Bartolucci and J. E. Besag (2002). "A recursive algorithm for Markov random fields." Biometrika, 89(3):724–730. A recursive estimation procedure for the full conditionals of MRFs as an alternative to the Brooks expansion. The technique uses ideas from time series factorisation and a block Gibbs sampler. As with Besag (1974), the factorisation is arbitrary and alternate factorisations must lead to the same joint distribution.

- J. E. Besag and D. Mondal (2005). "First-order intrinsic autoregressions and the De Wijs process." Biometrika, 92:909–920 Develops relationships between first-order intrinsic autoregressions and the de Wijs process which involves a continuous space differential equation.

More applied

- P. J. Diggle, J. E. Besag, and T. Gleaves (1976). "Analysis of spatial point patterns by means of distance methods." Biometrics, 32(3):659–667. Considers some distance based method to test for spatial randomness.

The emphasis is on their use at the preliminary stage of the investigation. T-square sampling is promoted as a quick and useful method that can be used between different populations and is suited to large populations.

- L. Knorr-Held and J. Besag (1998). "Modelling risk from a disease in time and space. Statistics in Medicine." 17:2045–2060, 1998. Develops ideas for modelling risk from a disease in time and space using the basic models introduced for image analysis.▲

SOFTWARE HIGHLIGHT

BAYESIAN TIME SERIES MODELS IN MCMCPACK

Jong Hee Park
jhp@uchicago.edu

Bayesian techniques have been widely applied in social science research, and there is little doubt that R packages for Bayesian inference [9] continue to play an important role in these applications. Among them, MCMCpack [6] supports Bayesian analysis for a wide range of statistical models, density functions, and pseudo-random number generators, as well as providing a number of utility functions for creating graphs, reading and writing data to external files, creating mcmc objects, and manipulating variance-covariance matrices. One characteristic that distinguishes MCMCpack from other Bayesian R packages is its integrated Bayesian approach to model-fitting and model diagnostics. In this article, I introduce some recent developments in MCMCpack, with a particular focus on Bayesian time series models.

There are two kinds of time series models in MCMCpack. First, there are change-point models (Markov mixture models or hidden Markov models) using discrete response data. MCMCpack uses Chib [2]'s efficient algorithm to sample hidden states from a variety of parametric models. For binary time series data, `MCMCbinaryChange` and `MCMCprobitChange` provide functions to find breaks and estimate regime-specific parameters given detected breaks. While regressors are not allowed in `MCMCbinaryChange`, users may employ the same modeling structure with a probit regression model in `MCMCprobitChange`. `MCMCoprobitChange` simulates from the posterior distribution of an ordinal probit regression model with multiple parameter breaks using the Cowles method [3]. The simulation of latent states is based on Park (2011) [8]'s method to avoid the problem of clustered ordered categories. Last, `MCMCpoissonChange` fits a Poisson regression model with multiple changepoints. To avoid computational difficulties in designing the proposal density for Metropolis-Hasting algorithms, `MCMCpoissonChange` utilizes Fruhwirth-Schnatter and Wagner (2006) [4]'s auxiliary mixture sampling method. The details of the model are discussed in Park (2010) [7]. For each of these change-point models, MCMCpack allows Bayesian model comparison using marginal likelihoods [1]. Thus, users can identify the "true" number of breaks by comparing several candidate models with varying numbers of breaks before proceeding to the interpretation of the posterior distributions of the "true" model.

In addition to the above-described change-point models, MCMCpack also includes dynamic models for two specialized social science model applications. First is the one-dimensional item response theory (IRT) model. `MCMCdynamicIRT1d` simulates from the posterior distribution using the algorithm of Martin and Quinn (2002) [5]. The model assumes that each subject has a given ability and that each item has both a difficulty parameter and a discrimination parameter. The observed choice is dictated by an unobserved utility which is a function of item-specific characteristics, subject-specific abilities, and disturbances. The second specialized dynamic model in MCMCpack is `MCMCdynamicEI`, which fits Quinn [9]'s dynamic ecological inference model for partially observed 2 x 2 contingency tables. `MCMCdynamicEI` incorporates a dynamic random walk prior into Wakefield (2001)[11]'s model of ecological inference to capture temporal dependence in 2 x 2 contin-

gency tables observed over time.

The functions included in MCMCpack have increased dramatically in recent years. However, thanks to its unified design, all of these functions share a standardized model interface, even as they cover a variety of model-specific, computationally efficient MCMC algorithms. MCMCpack continues to serve as an easy-to-use and highly reliable computational environment for a variety of statistical models in the social sciences. We, the developers of MCMCpack, welcome all comments and suggestions from users.

Acknowledgements

I thank the two other members of the the MCMCpack project, Andrew D. Martin and Kevin M. Quinn. I would also like to thank all the users of and contributors to MCMCpack including Michael Malecki, Daniel Pemstein, and Craig Reed.

References

[1] S. Chib. Marginal Likelihood From the Gibbs Output. Journal of the American Statistical Association, 90, 1313–1321, 1995.

[2] S. Chib. Estimation and Comparison of Multiple Change-Point Models. Journal of Econometrics, 86(2), 221–241, 1998.

[3] M. K. Cowles. Accelerating Monte Carlo Markov Chain Convergence for Cumulative-Link Generalized Linear Models. Statistics and Computing, 6, 101–110, 1996.

[4] S. Fruhwirth-Schnatter, Wagner H. Auxiliary Mixture Sampling for Parameter-driven Models of Time Series of Small Counts with Applications to State Space Modelling. Biometrika, 93, 827–841, 2006.

[5] A. D. Martin, K. M. Quinn. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999. Political Analysis, 10, 134–153, 2002.

[6] A. D. Martin, K. M. Quinn, J. H. Park. MCMCpack: Markov Chain Monte Carlo (MCMC) Package. R package version 1.0-11, URL http://CRAN.R-project.org/ package=MCMCpack.

[7] J. H. Park. Structural Change in U.S. PresidentsÕ Use of Force. American Journal of Political Science, 54(3), 766–782, 2010.

[8] J. H. Park. Changepoint Analysis of Binary and Ordinal Probit Models: An Application to Bank Rate Policy Under the Interwar Gold Standard Political Analysis, Forthcoming, 2011.

[9] J. H. Park, A. D. Martin, K. M. Quinn. CRAN Task View: Bayesian Inference. Version 2011-05-10, URL http://CRAN.R-project.org/view=Bayesian.

[10] K. Quinn. Ecological Inference in the Presence of Temporal Dependence. In Ecological Inference: New Methodological Strategies. Gary King, Ori Rosen, and Martin A. Tanner (eds.). New York: Cambridge University Press, 2004.

[11] J. C. Wakefield. Ecological Inference for 2 x 2 Tables. Journal of the Royal Statistical Society, Series A. 167(3): 385–445, 2004.▲

STUDENTS' CORNER

Luke Bornn
l.bornn@stat.ubc.ca

In this issue's Students' Corner, we continue our Q & A with a panel of leading Bayesian statisticians, including new member Paul Gustafson (University of British Columbia). If you have a question for the panel for future issues, please email me. Following the Q & A, find the dissertation abstract of Minghui Shi, entitled "Bayesian Sparse Learning for High Dimensional Data." If you are newly graduated and would like to publish your thesis abstract, don't hesitate to contact me.

Q & A

"In what circumstances would you recommend your students do a post-doc? In what circumstances is a post-doc not the best choice for graduates?"

Peter Green
p.j.green@bristol.ac.uk

This is another of those "it depends where you are" questions that Luke is so good at ... in this case, because the availability of post-doc positions varies from country to country, as do the implications for what happens next in your career. But in general, I would say "almost always" I would recommend a graduating student to do a post-doc.

However, you should do it for the right reasons. It is not about "buying time" to allow you to postpone a decision about your future, although this can of course be a useful by-product. You should think of it as a transition and an opportunity ... a transition between the relative security of being a graduate student, where duties are few and responsibilities for planning research are shared with your adviser, and the independence and greater demands of many kinds that immediately hit you in a faculty position. As a post-doc, you take a step in these directions while retaining support and mentoring, and without making long-term commitments. As an opportunity, it allows a period to try new directions in research, exploiting the freedom that comes from release from the straitjacket that many students feel constrains them, completing a thesis project to a deadline.

I think the post-doc experience works best and adds most if you take a step outside your comfort zone – consider a research project that is adventurous, and not perfectly aligned to your current experience and skills. And move to a different department. Too many post-docs who don't move spend their time with the same adviser and end up just doing the research they didn't have time for in their thesis.

Here in the UK, post-docs come in two flavours – research assistantships and research fellowships. In the first, the PI has been awarded funds to complete a specific project, and you are the human resource to help him or her complete it; in the second, you decide the project and aim to be largely self-directed – appointment depends not only on your past accomplishments, but also on your plans and their viability. The different models suits different people. There are not many fellowships, but a factor in choosing an assistantship is the degree of independence you will be allowed. Good luck!

Paul Gustafson
gustaf@stat.ubc.ca

My gut reaction is that doing a post-doc, whether by choice or necessity, is a good thing. Some time to focus on science, and transition from doctoral student to independent investigator without too many distractions, is a real opportunity to start a scientific career on a sound footing. This opinion is based both on my own experience of doing a one-year postdoc (albeit with some teaching added-in), and from what I've seen with others launching their careers. To facilitate the transition to independent investigator, I think a good postdoc supervisor is commensurately more hands-off than a Ph.D. supervisor, but more hands-on than a research collaborator.

Of course the feasibility of instant transition, straight from Ph.D. to tenure-track, varies not just with the graduate's CV, but also over time, as economic and demographic forces exert themselves. As well, national boundaries come into play. For instance, the Canadian granting system is currently such that research funding levels in the statistical sciences are too low to create very many postdoctoral positions, at least for purely methodological work. More interdisciplinary work may fare better.

As a supervisor, I've seen both an outgoing Ph.D. graduate and an incoming postdoc get the best of both worlds. Both accepted a tenure-track position straight out of the Ph.D, but then negotiated to start one year later, in order to do a post-doc year elsewhere in the meantime. This year is then particularly unencumbered and research-centric, with neither professorial duties or job-search duties at hand. Someone sufficiently talented and fortunate to realize this situation is giving themselves a particularly good opportunity to launch.

Dave Higdon
dhigdon@lanl.gov

A post-doc position is a great opportunity to focus on an application, work with new people,

publish in a new area, see someplace new and expand your experience, with only a two or three year commitment. The downside is that you've got a few more years of apartment living, and you'll likely have to move again when it's over. I know people who have done very well with and without taking a post-doc job – I doubt the post-doc was the deciding factor in their success.

From a professional standpoint, postdoc experience can be a big help. At Los Alamos we are always looking for people with experience and breadth, adept at connecting statistics to applications. In addition, a postdoc position affords the opportunity to build some expertise in a new application area. This is particularly important if your research is going to be guided by applications.

Of course, non-professional aspects of a new opportunity can sometimes outweigh the professional aspects. Taking a post-doc job is a big life decision with a number of important considerations like: Will you be excited about the work you'll be doing? Who will you be interacting with? Will you connect with these people? Will you have fun? Will this position/location work for you socially? Will your spouse be happy? Etc.

After 4+ years of grad school, you're finally in charge of what you'll be doing. Hopefully you know yourself well enough to decide if a particular postdoc will be a good fit. Lots of people will have lots of good advice and notions of what makes a good career. But in the end, it's your job to make your career into something you want. Don't forget to have fun along the way!

Fernando Quintana and Alejandro Jara
quintana@mat.puc.cl
ajara@mat.puc.cl

There is no one-size-fits-all answer here. On one hand, the answers to these questions lie on the student's interests and potentialities. On the other hand, not all post-docs are the same. It seems difficult to derive a general rule for such recommendations, so we limit ourselves to discussing a few relevant elements that can be helpful in this matter.

If the student is motivated by a research and/or academic career, then the recommendation is to definitively do a post-doc. Indeed, this would be a great opportunity to boost the research line started in the dissertation (or to try new or complementary ones), publish all the related papers (and start new ones!), and get a good glimpse at academic and research life from the insider's viewpoint. It is the opportunity of seeing oneself as a scholar. This is also the last opportunity to enjoy of lots of free time to focus on research work. Aside from these motivations, many academic positions require de facto a recommendation letter originating in a previous post-doctoral experience.

When there is no motivation for a research or academic career, a post-doc could still be a good way to acquire experience in some particular area of application. But if the student only wants a regular job or a teaching position, then the post-doc does not seem recommendable as it would only delay the final goal.

Christian Robert
xian@ceremade.dauphine.fr

Postdoctoral positions are an essential part of a researcher training that I recommend to all my PhD students, even though they do not always follow this recommendation! Especially in our French academic structure where (a) PhD students complete their PhD at 25 or 26, (b) the job market is rather favourable to statisticians, (c) regular university positions are tenured from the start, but (d) involve a heavy teaching load... Unless family issues are preventing one from spending one or two years abroad, the experience brought by spending this time in a foreign department with a different academic culture is almost invariably highly positive. And this for many reasons: a high level of freedom and time for conducting research, writing papers and fattening one's vita, no administrative burden, usually in a prestigious institution with a top quality research group, the opportunity to start collaborations with more senior researchers, most often no teaching, sometimes the opportunity of learning a new language, the possibility of discovering a new country, etc. Even though sabbatical years are available in most academic systems, this is clearly the freest of all times in one's life and taking the opportunity of a post-doc can shape one's academic and non-academic future! So, unless there are no postdoc positions available anywhere appealing (!), or those offered are in a topic that sounds too far from one's PhD research, or under conditions that are too constraining, there are few reasons to miss the benefits and the fun of doing a postdoc.

Stephen Walker
s.g.walker@kent.ac.uk

A PhD now is not a commitment to an academic career, but a post-doc surely is. Giving recommendations to someone about any career choice is fraught with difficulties and issues. My general attitude is not to suggest any career path to anyone (save my own children and parents), but to give full support to someone who is keen and knows what it is they want to do. But to answer the question. I would recommend my student do a post-doc if on some occasion I was asked "Look what I've done – what do you think?" And I see something I wished I had done myself or thought of myself. [Or some equivalent event to this]. And for the second part. When is a post-doc not appropriate? When the passion is not there. When the ability is not there. But the person who would know this the best is the student themself.

Mike West
mike@stat.duke.edu

Until the beginning of the 1990s, postdocs were almost unheard of in mainstream statistics (though more common in biostatistics). As statistics grew "out" into this modern interdisciplinary age, with more aggressive involvement of stochastic modelling and model-based statistical reasoning in larger, multi-faceted and longer term research programs, professional development evolved with substantial expansion of postdoc opportunities. Nowadays, graduating PhDs have opportunities to continue their research education, statistical and scientific exposure and broadening, and their early entree into research leadership and the profession with postdocs in a range of environments: from core statistics in academia, to interdisciplinary statistics in academia, government labs and industry, and in national institutes of various flavours. It is a continual delight for me – and should be for all of us – that our professional world is so robust and dynamic, and that the "sellers marker" for statisticians – especially Bayesians with strong applied and computational interests – is arguably unique in the range of exciting and rewarding early career opportunities; and I see nothing but continuation for decades to come.

To the question ... well, it depends on the individual of course. Professional maturity is the key; this operates on a different time scale for each of us while having multiple aspects that weight differently for each of us. This – to me – is the main point for consideration in looking at postdoc positions versus "real jobs" on graduation: the role of the PhD advisor here is to help students to honestly understand how "ready" they are to take on the challenges inherent in each of the opportunities they may have. Some students are obviously "ready" for a faculty or comparable non-academic position on graduation; students having had broader exposure to multiple research areas, interdisciplinary as well as "core" statistics projects, and a range of teaching experiences can and do move seamlessly into academic research/teaching positions, or positions with a comparable (broad!) spectrum of demands on intellect, energy and professional maturity in industry, government labs or elsewhere. Other students may simply be less broadly experienced and might more clearly benefit from what is typically a relatively relaxed and less demanding postdoc research position; postdocs can provide an immensely positive personal "time to breath" as well as the opportunities for professional and research growth. In other cases, a student may be quite ready for a "real job" but attracted to the opportunities a postdoc offers in new research directions, to broaden (or, in some cases, move away from!) thesis research areas, as well as the opportunities to work in a different research group/environment and with new senior mentors. Postdocs in statistics – whether "basic" research and in large-scale interdisciplinary environments – serve an increasingly active and substantial role here, and, for me, an immensely valuable one for the profession. Thankfully, the historical academic emphasis on moving new PhDs as fast as possible into prime academic positions – to the benefit of the PhD advisor and department "placement" record but sometimes not so obviously to the new PhD in the very early years – is less evident nowadays; broad ranges of postdoc opportunities are positive, desirable and increasingly important components of the professional infrastructure of statistical science, and I hope we have by now grown out of the historical view of postdocs as "second class" next-steps for new researchers.

Dissertation Abstracts

Bayesian Sparse Learning for High Dimensional Data

by Minghui Shi
ms193@stat.duke.edu
Department of Statistical Science,
Duke University
PhD Supervisor: David Dunson

In this thesis, we develop some Bayesian sparse learning methods for high dimensional data analysis. There are two important topics that are related to the idea of sparse learning – variable selection and factor analysis. We start with Bayesian variable selection problem in regression models. One challenge in Bayesian variable selection is to search the huge model space adequately, while identifying high posterior probability regions. In the past decades, the main focus has been on the use of Markov chain Monte Carlo (MCMC) algorithms for these purposes. In the first part of this thesis, instead of using MCMC, we propose a new computational approach based on sequential Monte Carlo (SMC), which we refer to as particle stochastic search (PSS). We illustrate PSS through applications to linear regression and probit models.

Besides the Bayesian stochastic search algorithms, there is a rich literature on shrinkage and variable selection methods for high dimensional regression and classification with vector-valued parameters, such as lasso and the relevance vector machine. Comparing with the Bayesian stochastic search algorithms, these methods does not account for model uncertainty but are more computationally efficient. In the second part of this thesis, we generalize this type of ideas to matrix valued parameters and focus on developing efficient variable selection method for multivariate regression. We propose a Bayesian shrinkage model (BSM) and an efficient algorithm for learning the associated parameters.

In the third part of this thesis, we focus on the topic of factor analysis which has been widely used in unsupervised learnings. One central problem in factor analysis is related to the determination of the number of latent factors. We propose some Bayesian model selection criteria for selecting the number of latent factors based on a graphical factor model. As it is illustrated in Chapter 4, our proposed method achieves good performance in correctly selecting the number of factors in several different settings. As for application, we implement the graphical factor model for several different purposes, such as covariance matrix estimation, latent factor regression and classification.▲

NEWS FROM THE WORLD

Call for Announcements

Sebastien Haneuse
haneuse.s@ghc.org

I would like to encourage those who have any announcements or would like to draw attention to an up-coming conference, to get in touch with me and I would be happy to place them here.

Announcements

2012 ISBA World Meeting

Planning has already begun for the 11th ISBA World Meeting, to be held in June 2012 in Kyoto, Japan. See the June 2010 issue of the ISBA Bulletin for the announcement and more information, as well as the conference website http://www2.e.u-tokyo.ac.jp/~isba2012/.

Meetings and conferences

III Latin American Meeting on Bayesian Statistics & XXXVIII National Meeting of Statistics, Pucón, Chile. 23-27th October, 2011.

The III COBAL is an international event whose purpose is to allow discussion and interaction among researchers, academics and students, with the aim to further strengthen scientific exchange within the Bayesian community in La-

tinamerica. This Congress has been performed twice. The first time was in 2002 in Ubatuba, Brazil, and the second time in 2005 in Los Cabos, Mexico. This year 2011, the University of Santiago de Chile is leading the organization of the event in collaboration with other sponsoring universities. The joint event will take place specifically in Pucón Headquarters at the University of La Frontera, University Partner in the organization, where the series of talks and conferences are scheduled. Spanish and English are the official languages.

Additional information can be found at http://cobal2011.usach.cl/.

8th World Congress in Probability and Statistics, Istanbul, Turkey. 9-14th July, 2012.

Jointly organized by the Bernoulli Society and the Institute of Mathematical Statistics and scheduled every four years, this meeting is a major worldwide event for statistics and probability, covering all its branches, including theoretical, methodological, applied and computational statistics and probability, and stochastic processes. It features the latest scientific developments in these fields.

The program will cover a wide range of topics in statistics and probability, presenting recent developments and the state of the art in a variety of modern research topics, with in-depth sessions on applications of these disciplines to other sciences, industrial innovation and society. It will feature several special plenary lectures presented by leading specialists. In addition, there will be many invited sessions highlighting topics of current research interests, as well as a large number of contributed sessions and posters.

The venue of the meeting is Grand Cevahir Hotel & Convention Center located in Istanbul which is a vibrant, multi-cultural and cosmopolitan city bridging Europe and Asia. Istanbul has a unique cultural conglomeration of east and west, offering many cultural and touristic attractions, such as Hagia Sophia, Sultanahmet, Topkapõ Palace and Maiden's Tower.

Additional information can be found at www.worldcong2012.org.

Short courses and workshops

Bayesian Inference in Stochastic Processes Workshop, Getafe (Madrid), Spain. 1-3rd September, 2011.

In this workshop, we will bring together experts in the field to review, discuss and explore directions of development of Bayesian Inference in Stochastic Processes and in the use of Stochastic Processes for Bayesian Inference. Theoretical and applied contributions to any area of Bayesian inference for stochastic process will be welcome. The workshop will be held in an informal environment to encourage discussion and promote further research in these fields. The workshop will be located in the Universidad Carlos III de Madrid, in the Getafe campus, less than a 20 minute train journey from the centre of Madrid.

Additional information can be found at http://www.est.uc3m.es/bisp7/.

Bayes-250 Workshop, Edinburgh, Scotland. 5-7th September, 2011.

The Schools of Mathematics and of Informatics at The University of Edinburgh are holding a research workshop to mark the 250th anniversary of the death of Thomas Bayes, a former student of the University of Edinburgh. The workshop will run from early afternoon on Monday 5th to late morning on Wednesday 7th September. The general theme of the workshop is that of what has come to be known as Bayesian statistics. Places at the workshop are limited, but there are still a number of places available.

For further information see the workshop web site http://conferences.inf.ed.ac.uk/bayes250.▲