THE ISBA BULLETIN

Vol. 10 No. 2 *August 2003*

The official bulletin of the International Society for Bayesian Analysis

Moving Forward and Going Strong: A Message from the President

by Ed George
ISBA President
edgeorge@wharton.upenn.edu

I am pleased to report that ISBA continues to move forward and is going strong. Let me begin with an exciting major announcement - the ISBA Board of Directors has formally approved production of a new electronic ISBA journal, Bayesian Analysis, with Rob Kass as founding Editor-in-Chief. Bayesian Analysis will be a new home for innovative research about Bayesian theory, methodology and application. It will serve as a highly visible centerpiece for ISBA, a magnet that will attract the very best Bayesian research and researchers worldwide. Rob Kass is currently assembling a stunning editorial board that will insure that the highest standards are established right from the outset. We are all extremely grateful to have someone so capable and of such tremendous stature as Rob to lead this pioneering effort. As I like to say, "What a good guy!" Stay tuned for further announcements as this new development moves forward.

As for going strong, numerous ISBA cosponsored meetings have already taken place around the world this year. I was fortunate to have already attended four of them - the Workshop on Objective Prior Methodology in Aussois, France, June 15-20, 2003, the first joint IMS-ISBA Conference in San Juan, Puerto Rico, July 24-26, 2003, Practical Bayesian Statistics 5, at the The Open University, U.K, July 28 - 31, 2003 and the International Workshop on Bayesian Data Analysis, Santa Cruz, CA, August 7-10, 2003. (Remember, information on these and future ISBA realated meetings can always be found on the ISBA website `http://www.bayesian.org/`).

To say only that these meetings were highly successful would be an understatement - they were awesome, both intellectually and socially! The high quality and tremendous breadth of Bayesian research was on display in full force at lectures and poster sessions. And in the Bayesian tradition of working hard and playing hard, interaction and fun were a central part of every meeting. But perhaps the most satisfying aspect of the meetings was the large influx of young new Bayesian researchers. This is a sure sign of the health and vitality of the Bayesian paradigm, and only reinforces my optimism for its continued future impact. All the organizers of these meetings, too numerous to mention here, did a spectacular job. Their generous hard work was deeply appreciated by all.

I should also add that all such ISBA cosponsored meetings serve to further increase ISBA membership. Last year the ISBA Board adopted a policy for ISBA cosponsored meetings that gives a registration discount for ISBA members, and gives a full year of ISBA membership to all attending non-ISBA members. In effect, this brings new members into ISBA, while providing a subsidy to the meetings. Please spread the word about this new win-win policy.

Turning to the ISBA awards, a hearty congratulations to all of this year's winners, and many thanks to all the prize committee members for their hard word. The Savage Awards for the outstanding doctoral dissertations in Bayesian econometrics and statistics went to Nicolas Chopin for Applications des Méthodes de Monte Carlo Séquentielles la Statistique Bayésienne and to Marc Suchard for Model Building and Selection in Bayesian Phylogenetic Reconstruction. The Mitchell Prize given in recognition of an outstanding paper that describes how a Bayesian analysis has solved an important applied problem was just awarded to Jeff Morris, Marina Vanucci, Phil Brown and Ray Carroll for Wavelet-Based Nonparametric Modeling of Hierarchical Function in Colon Carcinogenesis. It is interesting to note that this paper was also selected by the Journal of the American Statistical Association as their applications paper of the year. So the best Bayesian applications paper turns out to also be the best applications paper overall. Well, that makes perfect sense to me!

Finally, let me emphasize once more, that ISBA is a completely democratic organization. If you have ideas or suggestions for making ISBA better, please email them to me or any member of the ISBA board. We eagerly look forward to your involvement.

A MESSAGE FROM THE EDITOR

by Hedibert Lopes
hedibert@im.ufrj.br

Dear ISBA members, a whole paragraph was mistakenly included in the Student's corner of the last issue (March, 2003) of the Bulletin. I apologize for that. Also, the ISBA president and I have decided, for obvious reasons, that this issue will be an August Issue as opposed to the traditional June one. The third and final issue of 2003 will be the December issue. Hopefully next year we will return to the four-issues a year basis, and in that direction. I would like to take the opportunity to invite all members to contribute with the Bulletin. Please, feel free and encouraged to contact me or anyone from our Editorial Board at anytime. I like to think that the bulletin is the commonplace where one can get in touch with all Bayesians around the world. Enjoy the Issue!

REPORT ON THE THIRD WORKSHOP ON BAYESIAN INFERENCE IN STOCHASTIC PROCESSES

by Sonia Petrone
sonia.petrone@uni-bocconi.it

The third workshop on *Bayesian Inference in Stochastic Processes* took place in La Manga (Spain), from May 15th through May 17th, 2003. It followed the first workshop in Madrid (1998) and the second one in Varenna (Italy; 2001), and we are looking forward at the fourth workshop on this interesting area of research in Italy, in 2005. The workshop was a success, with about 50 participants from all over the word.

The talks and posters presented in the workshop gave an overview on recent research in Bayesian inference for problems where the data have a complex dependence structure, arising in different fields of application. So we had talks on spatial and/or temporal data, Poisson processes, point processes, diffusion processes, stochastic volatility financial series, empirical processes, processes in epidemics, queues models, survival analysis. Many papers treated the complex modelling issues, others had more emphasis on computational aspects, others were more focused on theoretical problems.

The workshop was held in an informal environment. La Manga is a beach resort near Cartagena, still very beautiful despite the enormous touristic development, situated on a tongue stretch of sand between a large laguna and the Mediterranean sea. The nice social dinner in the beautiful historic town of Cartagena confirmed once more the wonderful, warm and lively character of Spanish people.

The Scientific Committee (Armero, Cano, Gelfand, Kessler, Petrone, Rios Insua, Roberts, Ruggeri, Sanmartin, Soyer) thanks all the attendands for their lively scientific and social contribution to the success of the workshop, but a special thanks goes to Juan Antonio Cano, Mathieu Kessler and Pilar Sanmartin who did most of the work in efficiently organizing the workshop, in a splendid and warm atmosphere. Really muchas gracias to them!

Despite the wide interest for the applications, that was well represented and discussed in the workshop, I would say that it is evident that there is still the need of more theoretical work on Bayesian inference for stochastic processes. It seems to me that still great part of the theoretical research in Bayesian statistics deals with exchangeable processes, and one of the aims of these workshops is to stimulate more and more research on problems which involve more complex stochastic dependence structures.

So we look forward to seeing you all at the fourth workshop in Italy in 2005!

Contents

Opportunities for Bayesian Analysis in the Search for Sustainable Fisheries

by Marc Mangel and Stephan B. Munch
msmangel@ams.ucsc.edu and Zasummen@aol.com

There is general agreement among fishery scientists that many of the world's marine and freshwater fisheries have been overexploited and that many fish stocks are depleted and in need of rebuilding. There is also general agreement among scientists, the industry, the public and politicians that the search for sustainable fishing should receive high priority. Here, we explain how Bayesian analysis can make important contributions towards this goal.

To do this, we will use a simple heuristic model (at the end of the article, we direct interested readers towards appropriate literature). Imagine a fish stock with biomass at time t denoted by $N(t)$ which grows through reproduction and declines through natural and fishing mortality. One characterization of such dynamics is

$$\frac{dN}{dt} = \frac{\alpha N}{\beta + N} - MN - qEN \tag{1}$$

The first term on the right hand side of this equation is called the stock- recruitment relationship (this particular one is called the Beverton-Holt stock recruitment relationship, due to Raymond Beverton and Sidney Holt, who created modern quantitative fishery science just after world war II) and relates the current abundance $N(t)$ to the production of new individuals. At low population size the per unit biomass rate of reproduction is α/β. The parameter β captures the effect of density, and thus competition for resources, on the rate of reproduction. When abundance $N(t) = \beta$, the rate of reproduction drops to half of its maximum value. The parameter M is called the natural mortality rate and characterizes the rate at which biomass of the focal stock is lost due to parasites, predators and senescence. The combination qE is called the fishing mortality rate (often denoted by the letter F) and composed of a measure of fishing effort $E(t)$ and the catchability coefficient q, which is exactly what its name says: a measure of the ease with which fish are captured. Note that since catch, or yield, is qEN, catch per unit effort (CPUE) is a proxy for abundance. The time course of many fisheries is similar to that shown in Figure 1. Annual catch increases and then falls, cumulative catch continues to increase, and CPUE (the proxy for abundance) declines.

For a fishery to be sustainable, the right hand side of Eqn (1) must be non- negative. One may then begin by asking how the parameters determine the steady level of population size. Replacing qE by F, setting $dN/dt = 0$ and solving for the abundance at the steady state gives

$$\overline{N} = \frac{\alpha}{M+F} - \beta$$

so that the steady state yield is

$$\overline{Y} = F\left(\frac{\alpha}{M+F} - \beta\right).$$

It is a simple exercise in calculus to show that there is a value of F that maximizes yield (giving the maximum sustainable yield, MSY, which is now generally recognized as a constraint on harvest, rather than a target); this value is

$$F_{\text{MSY}} = \sqrt{\frac{\alpha M}{\beta}} - M$$

and corresponds to a steady state population size

$$N_{\text{MSY}} = \sqrt{\frac{\alpha\beta}{M}} - \beta.$$

These short calculations already show us an area in which Bayesian opportunities exist: the quantities of interest such as steady states and F_{MSY} involve functions of the parameters, so that any kind of maximum likelihood analysis of parameters from data such as that shown in Figure 1 will be confounded. Bayesian methods provide the natural means for going forward, especially since in many cases we are able to construct priors using biological knowledge about the species and information about fisheries for similar species elsewhere.

The stock-recruitment relationship can be written more generally as $R(N) = \alpha N f(N)$ where $f(N)$ is typically a nonlinear function taking values between 0 and 1, representing the decrease in the per capita reproductive rate occurring at high population densities. A number of alternative formulations for this function have been proposed. The most common of these is known as the Ricker stock-recruitment function in which $f(N) = exp[-\beta N]$. Although the derivations of the Beverton-Holt and Ricker models imply very different causal mechanisms underlying recruitment variability, in practice it is quite difficult to determine which model is appropriate.

The dynamics of the fishery may be quite sensitive to the choice of stock-recruit function and estimated parameters. For example, the equilibrium described above for a fishery whose population dynamics are governed by a Beverton-Holt stock-recruitment relation is always stable. In contrast, the equilibrium for a population governed by Ricker stock-recruitment relationship is stable only if $qE < \alpha - M$. Thus, determining which stock-recruitment function is appropriate is an important challenge, which may be addressed in a Bayesian framework both in terms of parameter estimation and in assigning posterior probabilities to particular choices of recruitment functions.

The dynamical behavior of the fishery at low population densities is also of critical importance. In this regard, a serious question is whether or not populations exhibit "depensation", which refers to the situation where at low populations sizes, recruitment is insufficient for population replacement. Depensation has been implicated in the collapse of several fish stocks including Cod in the Gulf of Maine. The question of whether a population exhibits depensation may be addressed by modifying the stock-recruitment relationship to allow an inflection point at low population size, say by the addition of another parameter γ, e.g. $f(N) = (1 + \beta N^{\gamma})^{-1}$ for Beverton-Holt or $f(N) = N^{\gamma-1}\exp[-\beta N]$ for Ricker. Here, an immediate opportunity for Bayesians concerns posterior inference for γ or perhaps more general inferences regarding the shape of the stock recruitment relationship at low population sizes.

The natural mortality rate M is an exceptionally difficult parameter to measure in nature. For that reason, fishery scientists have long sought proxies that could be related to mortality rate. One of these is individual growth rate k which is used to characterize the length of an individual of age A, $L(a)$ according to

$$\frac{dL}{da} = k(L_{\text{inf}} - L)$$

where L_{inf} is called the asymptotic size, since it is the maximum length that can be achieved (and is done so only as age becomes infinite). By appealing to evolutionary theory , one can show that under very general conditions there should be a positive relationship between k and M. The idea, then, is that one could then measure the growth rate of fish (or infer it from various calcified structures) and learn about the mortality rate. In consequence, we have nearly fifty years of data accumulated relating growth and mortality rates but only limited statistical approaches (mainly regressions) putting these data into a conceptual framework in which they could be used for predictive purposes.

The rate at which fish are removed by fishing is given by qEN. However, quantifying fishing effort is not straightforward. Fishing is carried out by varying vessels, using different fishing gears for disparate periods of time. Thus what constitutes a unit of fishing effort is subject to considerable uncertainty. Bayesian methods may be useful in designing robust effort indices. Moreover, the assertion that the rate of fishing is given by qEN is a simplification: the relationship between the number of fish captured and fishing effort may be expected to differ from simple proportionality for a number of reasons and could more generally be written as $Y = q(N)EN$. In this case the catchability $q(N)$ is some, potentially non- linear, function of biomass. Perhaps the most obvious reason for non- linearity in catchability is that fish school. Aggregation in schools means that large numbers of fish may continue to be captured even as the total number of fish in the population declines, potentially leading to unsustainable fishing. In such cases, CPUE is no longer a reasonable index of abundance and Bayesian inferences about how $q(N)$ changes with population size may be useful in avoiding overfishing.

In addition to natural mortality M and directed fishing mortality, the parameter F in Eqn (1), many fish stocks also experience incidental mortality, in which they are captured (and usually killed) in the course of a fishing operation not intended to take them. The determination of the size of such incidental mortality is another challenging problem, but beyond the scope of this brief review.

In managing a fishery, decisions must regularly be made regarding the total fishing effort that can be applied to the population or the total allowable catch. These decisions must be made in the presence of considerable uncertainty and must account for various social and economic issues as well. Bayesian decision analysis may be of great utility in weighing alternative management strategies in the presence of uncertain outcomes.

Getting into the literature. There is a long history of the applications of mathematics in fishery science. Smith provides an excellent description of activities until the mid 20th century.

Life history relationships between k and M were first introduced by Ray Beverton and Sidney Holt in the 1950s,in their studies of fishery management and their work on aging. They have been rediscovered many times (and probably will continue to be rediscovered).

In 1994, Beverton gave a series of lectures at laboratories of the National Marine Fisheries Service across the US. These were recently transcribed and published; they can be found at `spo.nwr.noaa.gov/BevertonLectures1994/`. Some other general sourcebooks are:

1. Hart, P. J. B., and J. D. Reynolds. 2002. Handbook of Fish Biology and Fisheries. Volume 2. Fisheries. Oxford, Blackwell.

2. Hilborn, R.and M. Mangel. 1997. The Ecological Detective. Confronting Models with Data. Princeton University Press.

3. Jennings, S., M. Kaiser, and J. D. Reynolds. 2001, Marine Fisheries Ecology. Oxford, Blackwell Science.

4. Quinn, T. J. I., and R. B. Deriso. 1999, Quantitative Fish Dynamics. New york, Oxford University Press.

5. Smith, T. D. 1994, Scaling Fisheries. Cambridge, UK, Cambridge University Press.

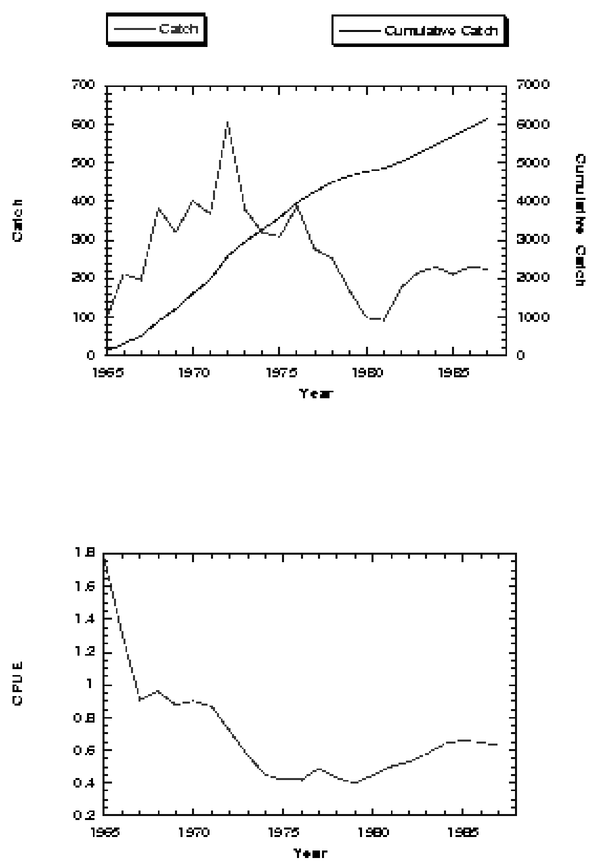

Figure 1: The annual and cumulative catch and catch per unit effort in the Namibian fishery for hake (from Hilborn and Mangel (1997)).

EVDBayes: BAYESIAN METHODOLOGY IN EXTREME VALUE THEORY

by Alec Stephenson
a.stephenson@lancaster.ac.uk

The role of extreme value theory is to develop procedures which are scientifically and statistically rational for estimating the extremal behavior of random variables or processes. Unique amongst all areas of statistics, extreme value theory is intrinsically about extrapolation; an extreme value analysis usually requires estimation of the probability of events that are more extreme than any that have already observed. Many purists argue that there can be no sensible rationale for performing such extrapolations, but there are many areas of application where they are necessary, so it is important to develop procedures that have some scientific basis. An introduction to the statistics of extreme values is given by Coles (2001) (*An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London)

The `evdbayes` package is an add-on package for the R statistical computing system, though most of the internal routines are implemented in the C programming language. It provides functions for the Bayesian analysis of extreme value distributions and models, using MCMC methods. The analysis is restricted to the likelihood models used most often in extreme value theory; those associated with maxima (or minima), those associated with upper (or lower) order statistics, and those associated with variables above (or below) a user specified threshold. Prior distributions can be constructed using either the multivariate normal distribution, or by constructing priors on the quantile space, for fixed probabilities, or on the probability space, for fixed quantiles.

It is still relatively rare for Bayesian methodology to be used within extreme value analysis. Many members of the extreme value community have little knowledge or experience with regard to Bayesian theory. This package has been produced in order to bridge this gap. The design of the package is therefore relatively simple, having only five main functions; one for each of the three methods of prior construction, one for optimizing (prior or posterior) target distributions, and one for generating Markov chains. Other functions exist for estimating features that are of particular interest in an extreme value analysis, such as return levels of predictive distributions.

The primary source of documentation for the package is The Users' Guide. This introduces Bayes theory and MCMC methods to extreme value theorists, and describes the extreme value likelihood models and methods of construction for prior distributions. Three data examples form the heart of the guide, illustrating the functions and the associated methodology. The `evdbayes` package and the documentation, including The Users' Guide, can be downloaded from `www.maths.lancs.ac.uk/~stephena`. Other R packages associated with extreme value theory can be downloaded from the same location.

FACTOR MODELS

by Hedibert Lopes
hedibert@im.ufrj.br

Factor models is certainly one of the most used (useful) statistical techniques. Factor models are mainly applied in two major situations: (i) data reduction and (ii) identifying underlying structures.

I would like to start this section by quoting Bartholomew (1995) (Spearman and the origin and development of factor analysis, *British Journal of Mathematical and Statistical Psychology, 48, 211-220*), who starts his paper's abstract by saying that

> Spearman [Charles Edward Spearman F.R.S. 1863-1945] invented factor analysis but his almost exclusive concern with the notion of a general factor prevented him from realizing its full potential.

Fortunately, factor models potentials have been discovered and are still being discovered, even after almost a century has passed since Spearman wrote his seminal paper ('General Inteligente' objectively determined and measured, *American Journal of Psychology*, 5, 201-293, 1904.)

I organized this annotated bibliography with the idea of providing the reader with a modest (and subjective) set of papers and books that would lead him/her to the realm of (latent) factor models.

Factor analysis: estimation

1. Lawley (1940) The estimation of factor loadings by the method of maximum likelihood, Proceedings of the Royal Society of Edinburgh, 60, 64-82. Lawley (1941) Further investigations in factor estimation, Proceedings of the Royal Society of Edinburgh, 61, 176-185. Introduces maximum likelihood factor model.

2. Jöreskog, K.G. (1967) Some contributions to maximum likelihood factor analysis, Psychometrika, 32, 443-382. Jöreskog, K.G. (1969) A general approach to confirmatory maximum likelihood factor analysis, Psychometrika, 34, 183-220. Maximum likelihood estimation is made feasible.

3. Martin, J.K. and McDonald, R.P. (1975) Bayesian estimation in unrestricted factor analysis: a treatment for Heywood cases, Psychometrika, 40, 505-517. Treatment of the Heywood case (zero variances) by proper specification of the prior distributions.

4. Geweke, J.F. and Singleton, K.J. (1980) Interpreting the likelihood ratio statistic in factor models when sample size is small, Journal of the American Statistical Association, 75, 133-137.
From Monte Carlo simulations and under certain regularity conditions, asymptotic theory is appropriate when sample size is greater than 30. The same is not true when the regularity conditions fail.

5. Bartholomew, D.J. (1981) Posterior analysis of the factor model, British Journal of Mathematical and Statistical Psychology, 34, 93-99. The posterior analysis is restricted to the common factors upon previous estimation of the model parameters (loadings and idiosyncrasies).

6. Lee, S-Y (1981) A Bayesian approach to confirmatory factor analysis, Psychometrika, 46, 153-160. A Newton-Raphson algorithm is implemented to find the posterior mode for four different prior specifications.

7. Rubin, D.B. and Thayer, D.T. (1982) EM algorithms for ML factor analysis, Psychometrika, 47, 69-76. Bentler, P.M. and Tanaka, J.S. (1983) Problems with EM algorithms for ML factor analysis, Psychometrika, 48, 247-251. Rubin, D.B. and Thayer D.T. (1983) More on EM for factor analysis, Psychometrika, 48, 253-257. The EM algorithm is introduced as an alternative optimization algorithm to Jöreskog's (1967,1969) maximum likelihood scheme.

8. Bartholomew, D.J. (1984) The foundations of factor analysis, Biometrika, 71, 221-232.

9. Bartholomew, D.J. (1985) Foundations of factor analysis: some practical implications, British Journal of Mathematical and Statistical Psychology, 38, 1-10. With discussion on the British Journal of Mathematical and Statistical Psychology, 38, 127-229.

10. Anderson, T.W. and Amemiya, Y. (1988) The asymptotic normal distribution of estimators in factor analysis under general conditions. The Annals of Statistics, 16, 759-771. Amemiya, Y. and Anderson, T.W. (1990) Asymptotic chi-square tests for a large class of factor analysis models. The Annals of Statistics, 18, 1453-1463. Asymptotic estimation and hypothesis testing in factor analysis models.

11. Press, S.J. and Shigemasu, K. (1989) Bayesian inference in factor analysis, in Contributions to Probability and Statistics: Essays in Honor of Ingram Olkin, L.J. Gleser, M.D. Perlman, S.J.Press, A.R.Sampson (Eds.), New York: Springer-Verlag, 271-287. Posterior large sample interval estimators of common factors, factor loading and idiosyncratic variances. Three step procedured: (1) estimation of the common factors, (2) estimation of the factor loadings given the common factors estimate, and (3) estimation of the idiosyncratic variances given both the common factors and factor loadings estimates.

12. Bartholomew, D.J. (1995) Spearman and the origin and development of factor analysis, British Journal of Mathematical and Statistical Psychology, 48, 211-220. Historical account of the development of factor models.

13. Ihara, M. and Kano, Y. (1995) Identifiabiliy of full, marginal, and conditional factor analysis models, Statistics and Probability Letters, 23, 343-350. Conditions for full, marginal and conditional model identification are discussed.

14. Schneeweiss, H. and Mathes, H. (1995) Factor analysis and principal components, Journal of multivariate analysis, 55, 105-124. Similarities and differences between the factor analysis and principal component analysis are discussed.

15. Yung, Y.-F. (1997) Finite mixtures in confirmatory factor-analysis models. Psychometrika, 62, 297-330. Finite mixture of factor models for handling heterogeneity. Approximate-Scoring (AS) and Expectation-Maximization (EM) methods are developed.

16. Lee, S.E. and Press, S.J. (1998) Robustness of Bayesian factor5analysis estimates, Communications in Statistics, Theory and Methods, 27, 1871-1893. Posterior robustness of the loadings, common factors and idiosyncratic covariance.

17. Fokoué, E. and Titterington, D.M. (2000) Bayesian sampling for mixtures of factor analysers, Technical Report, Department of Statistics, University of Glasgow. Gibbs sampler in mixture of factor models where the number of components and common factors are fixed and known.

18. Lopes, H.F. (2003) Expected posterior priors in factor analysis. (Brazilian Journal of Probability and Statistics, to appear), Technical report, Department of Statistical Methods, Federal University of Rio de Janeiro.

Factor analysis: model selection

19. Akaike, H. (1987) Factor analysis and AIC, Psychometrika, 52, 317-332. Model selection through the Akaike Information Criterion.

20. Press, S.J. and Shigemasu, K. (1994) Posterior distribution for the number of factors, in American Satistical Association Proceedings of the Section on Bayesian Statistical Science, 75-77. Large sample is assumed to approximate the factor model's predictive density, which is used for model comparison.

21. Polasek, W. (1997) Factor analysis and outliers: A Bayesian Approach. Discussion paper, University of Basel. The number of factors is determined by computing marginal likelihoods through Chib's (1995,JASA, 1313-1321) algorithm.

22. Bozdogan, H. and Shigemasu, K. (1998) Bayesian factor analysis model and choosing the number of factors using a new informational complexity criterion. Technical report, Department of Statistics, University of Tennessee.

23. Fokoué, E. and Titterington, D.M. (2000) Stochastic model selection for Bayesian mixtures of factor analysers, Technical Report, Department of Statistics, University of Glasgow. A birth-death marked Markov point process in continuous time is used as a stochastic model selection algorithm for Bayesian mixture of factor models.

24. Lopes, H.F. and West, M. (2003) Model assessment in factor analysis. Statistica Sinica (to appear) ISDS-Discussion Paper 98-38. A reversible jump Markov chain Monte Carlo (RJMCMC) algorithm is developed to fully account the uncertainty on the number of common factors. Comparisons are made to several additional algorithms that approximate the predictive density.

25. West, M. (2002) Bayesian Factor Regression Models in the "Large p, Small n" Paradigm. Discussion paper #02-12, ISDS, Duke University. Factor models where the number of variables is extremely larger than the number of observations, a situation commonly present in studies of gene expression.

Factor analysis in time series

26. Peña, D. and Box, G.E.P. (1987) Identifying a simplifying structure in time series. Journal of the American Statistical Association, 82, 836-843. Factor models with common (independent/dependent) factors following ARMA processes.

27. Engle, R. (1987) Multivariate ARCH with factor structures – cointegration in variance, University of California, San Diego, Dept. of Economics Discussion Paper 87-27. One of the first papers to apply common factors to model covariances in time series.

28. Diebold, F.X. and Nerlove, M. (1989) The dynamics of exchange rate volatility: a multivariate latent ARCH model, Journal of Applied Econometrics, 4, 1-21. Multivariate GARCH structures through one latent factor.

30. Engle, R.F., NG, V.K. and Rothschild, M. (1990) Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills, Journal of Econometrics, 45, 213-238. Factor-ARCH to model conditional covariance matrix of asset returns.

31. Ng, V., Engle, R.F. and Rothschild, M. (1992) A multi-dynamic factor model for stock returns. Journal of Econometrics, 52, 245-266. Relates dynamic and static factors to porfolio allocation in financial markets.

32. Lin, W-L. (1992) Alternative estimators for factor GARCH models - a Monte Carlo comparison, Journal of Applied Econometrics, 7, 259-279. Compares four frequentist estimators for factor GARCH models: two-stage univariate GARCH (2SUE), two-stage quasi-maximum likelihoo (2SML), quasi-maximum likelihood with known factor weights (RMLE) and quasi-maximum likelihood with unknown factor weights (MLE).

33. Molenaar, P.C.M. and Gooijer, J.G.D. and Schmitz, B. (1992) Dynamic factor analysis of nonstationary multivariate time series. Factor models with lagged common factors to account for the persistance in time series trends.

34. Bollerslev, T. and Engle, R.F. (1993) Common persistence in conditional variances, Econometrica, 61, 167-186. K-factor generalized autoregressive conditional heteroscedasticity (GARCH) models are discussed and conditions are given for covariance stationarity. They also study co-persistence in multivariate integrated GARCH models.

35. Harvey, A., Ruiz, E. and Shephard, N. (1994) Multivariate stochastic variance models, Review of Economic Studies, 61, 247-264. Common factors, as multivariate random walk, are used to model persistent movements in stochastic volatility models.

36. Escribano, A. and Peña, D. (1994) Cointegration and common factors. Journal of time series analysis, 15, 577-586. Cointegrated vectors are viewed as Peña and Box's (1987) dynamic factor models.

37. Geweke, J. and Zhou, G. (1996) Measuring the pricing error of the arbitrage pricing theory. The review of financial studies, 9, 557-587. First paper to implement the Gibbs sampler for exact Bayesian inference in (static) factor models.

38. Demos, A. and Sentana, E. (1998) An EM algorithm for conditionall heterescedastic factor models, 16, 357-361. Application of the EM algorithm to factor models with dynamic heterocedasticity in the common factors.

39. Sentana, E. (1998) The relation between conditionally heteroskedastic factor models and factor GARCH models, Econometrics Journal, 1, 1-9. Investigation of the similarities and differences of Engle's (1987) factor GARCH model and Diebold and Nerlove's (1989) latent factor ARCH model.

40. Pitt, M.K. and Shephard, N. (1999) Time varying covariances: A factor stochastic volatility approach (with discussion). In Bayesian statistics 6, Ed. Bernardo, J.M., Berger, J.O., Dawid, A.P. and Smith, A.F.M., 547–570. London: Oxford University Press. Univariate stochastic volatility structures is used to model the common factor variances through time. Sequential portfolio allocation is made possible by particle filters.

41. Aguilar, O. and West, M. (2000) Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. Journal of Business and Economic Statistics, 18, 338–357. Multivariate stochastic volatility structure is used to model the common factor variances through time.

42. Chib, S., Nardari, F. and Shephard, N (2003) Analysis of high dimensional multivariate stochastic volatility models. Technical Report, Nuffield College, University of Oxford. Feasible Bayesian inference (through MCMC algorithms) for multivarite stochastic volatility models in highly dimensional settings.
43. Vrontos, I.D., Dellaportas, P. and Politis, D.N. (2002) A full-factor multivariate GARCH model (2003). Classical and Bayesian estimation of a variant of the multivariate GARCH model with as many factors as variables. The order of variable problem is dealt with Bayesian model averaging.
44. Fiorentini, G., Sentana, E. and Shephard, N. (2003) Likelihood-based estimation of latent generalised ARCH structures. Technical Report, Nuffield College, University of Oxford. Fast MCMC and simulated EM algorithms for latent factor GARCH model.

Books

42. Lawley, D.N. an Maxwell, A.E. (1971) Factor analysis as a statistical method (2nd edition). London: Butterworths.

43. Anderson, T.W. (1984) An introduction to multivariate statistical analysis (2nd edition). New York: John Wiley & Sons, Inc.

44. Krzanowski, W.J. and Marriott, F.H.C. (1995) Multivariate analysis - Part 2: classification, covariance structures and repeated measurements. Kendall's Library of Statistics 2. London: Arnold.

45. Bartholomew, D.J. and Knott, M. (1999) Latent variable models and factor analysis (2nd edition). Kendall's Library of Statistics 7. London: Arnold.

46. Basilevsky, A. (19??) Statistical factor analysis and related methods: theory and applications. New York: John Wiley & Sons, Inc.

47. Press, S.J. (1982) Applied multivariate analysis: using Bayesian and frequentist methods of inference (2nd edition). New York: Krieger

48. Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) Multivariate Analysis. London: Academic Press.

STUDENT'S CORNER

by Lilla Di Scala and Luca La Rocca
lilla@dimat.unipv.it luca@dimat.unipv.it

Dear Members of ISBA, we would like to use this Student's Corner as an opportunity to send out a request to all of you. As some may already be aware of, we are running a survey on the following topic:

"Which groundbreaking/essential papers do you believe that a graduate student with a serious interest in Bayesian statistics should not miss reading?"

In other words, if you had to decide on next years' choices for a Journal Club, which Bayesian papers would you include? We would very much appreciate if you could spare five minutes of your time and make a list of 3 or 4 such papers, whether methodological or computational, whether your own or the product of other distinguished authors. Please keep in mind that it would be very useful for students to have such a hit-list of papers which are considered a must from those within the field. We welcome all replies and hope these will be copious, looking forward to featuring the results in the next issues of the Bulletin. Finally, a note of thanks goes to all of you who have already responded!

➤ Phd Thesis

Let us now introduce this issue's two doctorate thesis abstracts.

The first one is from the Department of Mathematics at the University of Pavia, Italy. The PhD in Mathematics and Statistics (which actually foresees two distinct curricula) is a three-year program with an extensive course offering during its first half. The Mathematical Statistics graduate program is relatively new and is currently going into its seventh year. On the other hand, the University of Pavia is one of the oldest places of study in Italy. The town very much resembles an over-seas campus as the university and its population of students are the center of most of the town's activity.

Igor Pruenster - igor@eco.unipv.it
Department of Political Economics and Quantitative Methods University of Pavia, Italy
Thesis title: *Random probability measures derived from increasing additive processes and their application to Bayesian statistics*
Advisor: Eugenio Regazzini.

Abstract

Increasing additive processes (IAP), i.e. processes with positive independent increments, represent a natural tool for defining random probability measures whose distributions act as nonparametric priors for Bayesian inference. It is well-known that the celebrated Dirichlet process can be obtained either by normalizing a time-changed gamma process or, as a particular case of neutral to the right (NTR) random probability measure, by the exponential transformation of a suitable IAP. A new class of random probability measures is introduced by generalizing the former construction to any IAP: a normalized random measure with independent increments (RMI) is defined by a suitable normalization of a time-changed IAP. Even if their finite-dimensional distributions are not generally known, quantities of statistical interest turn out to have, for a large subclass, appealing forms leading to simple rules for prior specification and to predictive distributions which consist of a linear combination of the marginal distribution and of a weighted empirical distribution. Particular attention is devoted to the study of their means. Necessary and sufficient conditions for finiteness together with their exact prior and approximate posterior distributions are provided. Some illustrative examples of statistical relevance are considered in detail. Normalized RMI are then further generalized to normalized IAP driven random measures, which contain the popular mixture of Dirichlet process as a particular case. Conditions for their existence are given. In particular, results for the distribution of means under both prior and posterior conditions are derived, and, relying upon the introduction of strategic latent variables, a full Bayesian analysis is undertaken. Finally NTR random probability measures are considered. Their means can be represented as the exponential functional of an IAP. This fact is exploited in order to give sufficient conditions for finiteness of the mean and for absolute continuity of its distribution. Moreover, expressions for its moments, of any order, are provided. By resorting to the maximum entropy algorithm, an approximation to the density of the mean of a NTR prior is obtained. The numerical results are compared to those yielded by some well-established simulation algorithms.

Our second abstract is from a doctorate thesis obtained within the Program in Applied and Computational Mathematics at Princeton University. Miss Papavasiliou has told us that "(As) it is a program, the Faculty is also associated with other departments, usually either Mathematics, Engineering or Economics.

This gives students great flexibility in choosing their field. A student interested in Probability and Statistics usually works with Faculty members at the Operations Research and Financial Engineering or Mathematics departments. The program is not big (there are around 24 student) and it is very friendly. More generally, Princeton University is a very stimulating place academic-wise." The University is located in Princeton, New Jersey, which is a beautiful and quiet little town.

Anastasia Papavasiliou - pp2102@columbia.edu
Department of Applied Physics & Applied Mathematics Columbia University, USA
Thesis title: *Adaptive Particle Filters with Applications*
Advisor: Rene Carmona.

Abstract

In stochastic filtering, the goal is to compute the conditional distribution of a Markov chain, at any time point, given some partial information on it ('optimal filter'). We study the case where the transition kernel of the Markov chain depends on some unknown parameters. Following a standard Bayesian technique, we treat the parameters as a non-dynamic component of the Markov chain. If the kernel is mixing for every fixed value of the parameter and the parameters satisfy some identifiability assumption, we show that the posterior distribution will converge to a delta function on the true value, for any reasonable prior. Then, if the model is sufficiently continuous with respect to the parameters, the optimal filter of the augmented system will eventually converge to the optimal filter of the correct model. The optimal filter is computed by using some variation of the Interactive Particle Filter. This method allows for the online estimation of the parameters in nonliner filtering and has many applications. We present two of them. One is on estimating the stochastic volatility of a stock. The second is on separating two Markovian signals, when we only observe a linear combination of them.

NEWS FROM THE WORLD

by Gabriel Huerta
ghuerta@stat.unm.edu

✻ *denotes an ISBA activity*

➤ Events

International BCB-Workshop on Statistics and Cancer Genomics *August 21, 2003 at Magnus-Haus, Am Kupfergraben 7, 10117 Berlin*

Invited Speakers Edgar Brunner, (Universitat Gottingen) Peter Buhlmann, (ETH Zurich) Christian Hagemeier, (Charite Berlin) Michael Hummel, (Benjamin Franklin, Berlin) Joe Nevins, (Duke medical center) Christine Sers, (Charite, Berlin) Terry Speed, (UC Berkeley) Mike West, (Duke University)

Scope and Aim of the workshop: The analysis of microarray experiments in cancer genomics requires the cooperation of scientists from medicine, pathology, molecular biology, bioinformatics and statistics. The international BCB-Workshop on statistics and cancer genomics wants to bring these people together. The key questions are: What kind of statistics is needed in modern genomics, and what can statistics provide ?

Organizer: Rainer Spang (BCB / MPIMG)

Max Planck Institute for Molecular Genetics Ihnestr. 63-73, D-14195 Berlin phone: +49/30/8413 -1175, fax: -1176 email: spang@molgen.mpg.de http://compdiag.molgen.mpg.de/ibcb/

✻ **A Conference in Honor of Arnold Zellner: Recent Developments in the Theory, Method, and Application of Information and Entropy Econometrics** *September 19-21, 2003, Washington, DC*

Conference Objectives: This conference will (i) study and explore Information Theory (IT) solutions for linear estimation and inference problems, (ii) facilitate the exchange of research ideas in the field, (iii) promote collaboration among researchers from different disciplines, and (iv) highlight the major trends in Information and Entropy Econometrics.

In particular, this Conference will concentrate on the most recent (theoretical and applied) research in linear Bayesian and non- Bayesian, IT and entropic procedures with emphasis on modeling and measuring information. In addition, the conference will deal with the interpretations and meaning of the solutions to IT estimation and inference (e.g., statistical meaning, complexity and efficiency as well as informational meaning). Both theory and innovative applied papers will be included.

Conference Topics: Conference topics include theory and methods and applications in various fields. Theory and methods topics include: Measuring information Interpretation and meaning of IT solutions to estimation and inference Information in moment based estimation Bayesian analysis and IT Linear IT and Entropic procedures Optimal information processing

Applications topics from all fields and interdisciplinary papers within and between the following fields are particularly welcome: Economics and social science Business and management science Natural and physical science Medical and biological science

Structure: 1.Paper presentations (sessions consisting of 3 papers and discussants). 2.Invited Lecturers. 3.Invited Survey-Review Lectures.

More information at: http://www.american.edu/cas/econ/faculty/golan/conference.htm

✻ **ISBA 2004 World Meeting** *May 23-27, 2004, Chile*

It is a great pleasure to announce that the ISBA 2004 World Meeting will be held in Chile in May of 2004. It is scheduled from 23-27 May 2004 in Via Del Mar at the Hotel Del Mar Convention Center.

The Local Organizing Committee is chaired by Pilar Iglesias, Pontifica U. Catolica de Chile (pliz@mat.pu.cl) and Fernando Quintana, also Pontifica U. Catolica de Chile (Quintana@mat.puc.cl). The Scientific Committee is chaired by Fabrizio Ruggieri of CNR IMATI (fabrizio@mi.imati.cnr.it). The Finance Committee is chaired by Alicia Carriquiry from Iowa State University (alicia@iastate.edu).

Please plan to join us for a major international Bayesian gathering that will maintain the tradition of providing an exciting scientific environment along with a glorious physical environment. More details will emerge over the next few months and will be made available on the ISBA website: www.bayesian.org.

Applied Bayesian Statistics School GENOMICS: METHODS AND COMPUTATIONS *June 15-19 2004, Trento, Italy*

Preliminary announcement

CNR-IMATI (Istituto di Matematica Applicata e Tecnologie Informatiche at Consiglio Nazionale delle Ricerche) and the University of Pavia (DEPMQ) are planning to organise every year a School on state-of-the-art Bayesian applications, inviting leading experts in the field.

The topic chosen for the 2004 School is "Genomics: methods and computations" and the lecturer will be Mike West, the Arts & Sciences Professor of Statistics and Decision Science at the Institute of Statistics and Decision Sciences, Duke University, Durham, NC, USA. His web page is: `www.stat.duke.edu/ mw`

Professor West is Co-Director of the Computational and Applied Genomics Program (CAGP) at the Institute for Genome Sciences and Policy, Duke University (see `www.cagp.duke.edu/index.shtml`)

The Computational and Applied Genomics Program is a multidisciplinary research program concerned with the creation, modeling, analysis and integration of multiple forms of data in basic and clinical biomedical studies. The core foci of research in CAGP projects lie in basic genome science, complex modeling and statistical methods, and computational biology.

CAGP projects include both basic and applied research, with the applications arising from several related, multidisciplinary projects in specific biomedical areas. The applied contexts currently include projects involving molecular characterization studies, gene pathway studies and biological information integration in breast cancer and other cancers, cardiovascular studies, and neurological studies. Emerging projects include applications in environmental health areas.

The School will have at most 40 participants. Discounted fees will be offered to researchers from Universities and Research Centers. Special fees will be available to Ph.D students. Reasonably priced accommodation is available at the Congress center. Details on registration, room reservation, fees, and payment, will be provided with the first announcement, due in early Autumn, 2003.

Trento can be easily accessed by train and motorway. The airports of Verona, Venezia and Bolzano are quite close to Trento. Trento is in the northern part of Italy, very close to the Alps (Dolomiti).

People interested in attending the School can contact Professor Fabrizio Ruggeri at `fabrizio@mi.imati.cnr.it`. The School website is `www.mi.imati.cnr.it/conferences/abs04.html`.

✻ IMS-ASA's SRMS JOINT MINI MEETING ON THE CURRENT TRENDS IN SAMPLE SURVEYS AND OFFICIAL STATISTICS *January 2 - 3, 2004, Calcutta, India*.

Registration deadline: July 31, 2003

Sponsors: Institute of Mathematical Statistics (IMS), Survey Research Methods Section (SRMS) of the American Statistical Association.

Co-sponsors: The U.S. Census Bureau,The Gallup Research Center,Department of Statistics, Calcutta University.

Meeting webpage: `http://www.jpsm.umd.edu/ims`.

2003 NBER/NSF Time Series Conference In Honor of George Tiao's Retirement *September 19-20, 2003, Chicago*.

The meetings will be held at the University of Chicago Graduate School of Business Gleacher Center located at 450 N. Cityfront Plaza Drive in Chicago.

Lodging: Reservations must be made prior to August 28 in order to obtain this discount rate. There are a limited number of rooms available so reserve early. Contact information for Hyatt Regency Chicago (800 233-1234), 151 East Wacker Drive Chicago, Illinois 60601 USA.

Statistica Sinica will publish a special issue in honor of George Tiao's retirement titled "Bayesian inference, environment statistics, time series analysis, and their applications". The deadline for submission is November 15, 2003 (about one month after the conference). All submissions should be sent to the Statistica Sinica editorial office as the regular manuscript, but with a cover letter mentioning the special issue in George's honor. Interested authors are encouraged to submit electronically at `http://www.stat.sinica.edu.tw/statistica`. The issue is targeted to be published at the end of 2004. All submissions will undergo the usual refereeing process.

Current Thinking on The Use of Bayesian Analysis in the Pharmaceutical Industry in Drug Discovery, Regulatory Approval and Economic Evaluation, *Tuesday, 14th October 2003, London, UK and Wednesday, 22nd October 2003, Washington DC, USA*.

Henry Stewart are pleased to announce the following European and US events. This is a detailed, practical briefing on the use of Bayesian analysis in the pharmaceutical industry, how to get it right and in particular when it is the preferred way to go.

For all Biostatisticians, Statisticians, Biometricians, Health Economists, Epidemiologists and all those using Bayesian statistics in the pharmaceutical industry. All ISBA members are entitles to a 10% discount.

Full documentation will be provided to all delegates and adequate time set aside for a full questioning of the speakers. Full details can be found on the Henry Stewart website `www.henrystewart.com` or alternatively by contacting Dr Carlos Horkan, Tel: +44 207 404 3040; Fax: +44 207 404 2081, Email: `carlosh@henrystewart.co.uk`.

SUGGESTIONS

PLEASE, FEEL COMPLETELY FREE TO SEND US SUGGESTIONS THAT MIGHT IMPROVE THE QUALITY OF THE BULLETIN

`hedibert@im.ufrj.br`