Editors' introduction

Luc Bauwens[a], Wolfgang Polasek[b], Herman K. van Dijk[c]

[a]CORE, Louvain-la Neuve, Belgium
[b]University of Basel, Basel, Switzerland
[c]Tinbergen Institute and Econometric Institute, Erasmus University, Rotterdam, The Netherlands

The purpose of the recently established International Society for Bayesian Analysis (ISBA)[1] is:

To foster the study of Bayesian statistical and decision analysis, its relation to other methodologies, and its applications to scientific, industrial, and social problems.

The proceedings of the first Riverboat Conference, which was held from April 29 until May 4, 1993, on a cruiseboat going from Basel to Amsterdam on the river Rhine, fit the purpose of the above mentioned constitution remarkably well. The conference dealt with an important part of the historical analysis of statistics, in particular, the contributions of the Bernoullis of Basel; with some methodological issues in modeling and in the computation of probabilities and with applications of Bayesian statistical analysis to such scientific fields as economics, finance, and geology.

One may argue that the main ideas in Bayesian analysis refer to reporting and communication of probabilities and to efficient methods for the computation of these probabilities. Careful reporting of ideas which are present before (or *prior* to) the analysis and careful presentation of (*posterior*) results after an analysis give Bayesian studies of scientific, industrial, and social problems credibility. With respect to the computational aspect it was once (at one of the Valencia meetings in Bayesian statistics) revealed to the third author of this introduction that Bayesians have three commandments:

Integrate   ...   Integrate   ...   Integrate

[1] Article 1 of the draft constitution of ISBA.

In recent decades, these commandments have induced many researchers to work in the area of simulation techniques where sophisticated numerical integration methods have been developed which make use of Monte Carlo integration principles; see, e.g., Van Dijk et al. (1995) and the references cited here. The mathematical operation of integration can be simplified to the operations: summation and averaging. Summation and averaging of prior and sample information is the key of the Bayesian machinery.

The papers in this volume, which are only a partial selection of the papers presented at the conference, deal with historical, modeling, and computational issues of probabilistic inference, and deal with several applications in the fields of economics, finance, and geology.

In the first paper, Stephen Stigler considers briefly the contribution of the Bernoullis of Basel to statistical inference. Jakob (Jaque), Johann (Jean), and Daniel Bernoulli are the big three 'Bernoulli scientists' of Basel. Jakob and Johann were brothers and contributed substantially to the development of probability and statistics around the year 1700; Daniel Bernoulli (the nephew of the above) made his contribution in the second half of the eighteenth century. Daniel used in 1769 a method which can be phrased in today's language of robust statistics as an iterative M-estimate. It might be a surprise, but Stigler claims that D. Bernoulli had the implicit notion of a likelihood function which has to be maximized. D. Bernoulli multiplied together semicircles as a weighting function back in 1778, but he could only handle three observations. There was a famous discussant for his paper: Leonhard Euler, another son of Basel.

Shafer discusses Jacob Bernoulli's contribution to the philosophy of probability as indicated in his *Ars conjectandi*, for the interpretation and application of probabilistic inference today. The interaction between applications and the theory of probability was in the seventeenth century not formulated in a very precise way. Shafer discusses the historical development of the concept of probability as presented by Laplace, the frequentists, and twentieth century Bayesians. Shafer argues that the Frequentist–Bayes dispute over the meaning of probability is getting out of fashion since new areas of applications, like expert systems and probabilistic forecasting, require a theory of conjecture instead of a theory of probability.

Positive economics is a long disputed issue in economics but has never been discussed thoroughly from a subjective probability point of view. Milton Friedman's 1953 book still challenges the profession, and Pelloni discusses 'those features of Friedman's view which are essential for its neo-Bayesian interpretation'. He argues that, in order to be consistent with his subjectivism, Friedman is implicitly subscribing to the De Finetti and Savage view that induction amounts to 'how a set of rational beliefs should be modified in the light of new information'.

All modeling is the formulation of prior information, be it as exact restrictions or as stochastic restrictions. In formulating models certain rules are used. For

this purpose Zellner states: 'just as with the derivation of models, it would be desirable to have a dependable, reproducible method for formulating diffuse and informative priors'. In his paper many standard and other econometric models are derived using a formal, reproducible, maxent approach. A direct link is made between the maximal data information prior (MDIP) approach, the Jeffreys' priors, and entropic priors. Through the interaction of theoretical principles and examples the MDIP prior approach is presented for many stochastic models.

The next two papers deal with Markov Chain Monte Carlo methods. First, Liu and Rubin discuss a procedure to speed up the convergence of Markov Chain Monte Carlo methods. The usual approach is to simulate in an iterative way from a sequence of transition distributions (usually of the same form), starting from an initial value distribution, such that the target distribution is the stationary distribution of a Markov process. Their basic idea is to use a set of series, each independently drawn from the same distribution, and estimate moments of the target distribution efficiently using maximum likelihood. The estimated moments are then used in a restarted Markov Chain approach. In a situation where massive parallel computing is possible, this procedure may lead to substantial gains in computational efficiency.

Next, Chib discusses a Bayesian procedure using Gibbs sampling to estimate the parameters of a Markov Mixture Model (MMM). The key feature of an MMM is that the components of the mixture are for each observation selected according to an unobserved Markov process. As a consequence the likelihood function is, in most cases, not known in terms of elementary functions. Using Markov Chain simulation one can avoid the explicit evaluation of the likelihood by treating the unobserved data as additional unknown parameters and by simulating these together with the other parameters of the model. Chib shows how to generate from the joint distribution instead of generating from several conditional distributions. Further, a stochastic version of the EM algorithm is compared with a full Bayesian approach. Applications involve Gaussian and non-Gaussian data and autoregressive time series of US gross national product.

Whether economic time series can be characterized as stationary around a deterministic trend or as random walks (with drift) is an issue which has received much attention of economists. The implications of the two models are vastly different. For instance, Schotman and Van Dijk (1991), using annual data on US real gross national product over a long period, show that the random walk with drift can explain much of the variation of the data such that the remaining variation behaves like white noise and it appears that cyclical variation has disappeared. A random walk model corresponds to a unit root in the autoregressive representation. Bayesian analysis of the unit root centers around the choice of the prior; see, e.g., Phillips (1991) and the references cited there. Kadane, Chan, and Wolfson discuss the choice of the prior from the subjective

perspective and suggest an elicitation procedure to specify parameters of a piece-wise normal inverted-gamma family which is closed under sampling. It is of considerable interest to investigate whether the authors' approach can be generalized to the case of cointegrated series.

Young and Pettit introduce a new measure to determine the influence of a prior distribution on a Bayes factor. This measure is based on the idea of comparing two Bayes factors under different prior distributions, where one is an information prior while the other one is a diffuse prior. Demonstrating the approach for the linear model, the proper prior is chosen to be from the Normal-gamma class while the diffuse prior is based on three choices leading to three Bayes factors which are known as (i) the imaginary sample, (ii) the fractional, and (iii) the intrinsic Bayes factor.

Geweke proposes procedures for Bayesian inference in the reduced rank regression model. Examples of reduced rank models are the incomplete simultaneous equation model, the cointegrated vector autoregressive model, and factor models. He provides a Gibbs sampling algorithm for computing posterior moments under a known rank value and for computing predictive odds ratios for different rank values. The procedure is applied to find the number of factors in an asset pricing model.

Multivariate nonstationary time series are modelled by Kato, Naniwa, and Ishiguro by a state space representation which decomposes the vector time series into a trend, a cyclical autoregressive component, a seasonal, and an irregular component. The estimation procedures uses the Kalman filter and the AIC criterion, while the interpretation follows the so-called power contribution analysis of Akaike (1967) which is obtained via the spectral representation involving the industrial production and the wholesale price index of Japan. The paper shows an alternative for modelling nonstationary vector time series without pre-adjustments of the nonstationarity in the data.

Poirier provides a Bayesian analysis of estimation, testing and prediction in nested multinomial logit models. He suggests a class of prior distributions which combines conveniently with the likelihood and which depends on hyper-parameters over which researchers may have prior beliefs. He argues that such prior beliefs are typically implicit in non-Bayesian inference methods applied to this class of models.

Schotman implements a Bayesian approach to estimate the parameters of the Vasicek model of the term structure of interest rates. He uses times-series data on the short-term interest rate to build a prior on the parameters, and combines it with the information that can be extracted from cross-sectional data on the term structure at a given date. The cross-sectional data is subject to extreme collinearity, which motivates the use of extraneous information. An application to a dataset of Dutch bond prices is included.

West studies Bayesian inference in successive sampling discovery models. The context is that of sampling without replacement from a finite population (as for

example, when oil pools are successively discovered in a given area). The purpose of the statistical analysis is to make inferences about characteristics of the finite population, such as predicting summaries of the remaining units. West uses a superpopulation structure that requires to introduce latent quantities that simplify the unequal probability sampling scheme. Then a Gibbs sampling algorithm can be applied to make the inferences. The procedure is applied to a set of data on oil discovery which has been analysed previously by maximum likelihood techniques.

The papers in this volume show the continued progress in Bayesian reporting, Bayesian computing and Bayesian model choice. Interpreting Shafer broadly. progress in science occurs most fruitfully when new models and methods are applied to many cases. The different applications treated in this volume are a witness of the richness and variation of applications of Bayesian analysis. Thus a small step is made towards fulfilling the purpose of the International Society for Bayesian Analysis.

References

Akaike, Hirotugu, 1976, Canonical correlation analysis of time series and the use of an information criterion, in: R.K. Mehra and C.G. Lainiotis, eds., System identification: Advances and case studies (Academic Press, New York, NY) 27–96.

Phillips, Peter C.B., 1991, To criticize the critics: An objective Bayesian analysis of stochastic trends, Journal of Applied Econometrics 6, 333–364.

Schotman, Peter and Herman K. van Dijk, 1991, On Bayesian routes to unit roots, Journal of Applied Econometrics 6, 387–401.

Van Dijk, Herman K., Alain Monfort, and Bryan W. Brown, 1995, Econometric inference using simulation techniques (Wiley, New York, NY).