Teaching Bayesian Statistics : From Economic to Statistical Rationality

J P Florens - Toulouse, France
Michel Mouchart - Louvain-la-Neuve, Belgium

1. Introduction and summary

In this paper we examine, from the point of view of teaching statistical methods to future economists, some relationships between the concepts of rationality in economics and in statistics. Based on the authors' research and teaching experience, this paper discusses both the major benefits and difficulties in teaching statistical methods for economists following a Bayesian stream of thinking.

In the first part, we examine reasons for believing that Bayesian reasoning may be useful in the statistical education of future economists. The Bayesian approach may be viewed first of all as a theory about learning that is very natural when modelling economic behaviour under uncertainty. Thus, the statistician's "learning by observing" and the economist's "learning with a purpose (decision-making)" are most easily made congruent within a Bayesian framework. Teaching Bayesian thinking therefore endows the student with a better understanding of economic models and with statistical tools appropriate for the theories he will confront with observation. Furthermore, statistical methods in economics and in management are most naturally used within a decision theory framework for which the Bayesian approach is particularly relevant.

Also, students are easily convinced that specifying operational models requires specifying distributions on imperfectly known parameters: expert opinions (think, for instance, of the "tea room elasticities" used in applied general equilibrium models), unobservable individual heterogeneity components, and deficiencies in available economic data are examples of uncertainties which are most easily represented in terms of prior probabilities. Insofar as economic theory typically does not provide the statistician with a completely specified model, the Bayesian framework appears to be well-suited to discussing the concept of a "best" model, thanks to the concept of encompassing. Finally, presenting the results of a statistical analysis in a Bayesian format appears to be easier for students in economics to understand than in a classical (or sampling theory) format.

The second part of this paper discusses some difficulties which should not be underestimated when considering this pedagogical approach. Bayesian thinking systematically relies on probabilistic arguments, such as marginalising or conditioning, that seem elementary, but that students in economics generally do not master. Indeed, a Bayesian approach helps students become familiar with a form of reasoning that underpins much modern economic thinking. Another difficulty concerns the availability of computer programs that are sufficiently user-friendly and widely applicable. There is no doubt that the situation is improving every day but at present it is crucially important that the teacher have available programs which allow students to analyse actual economic data with the same ease as is possible for classical methods. Keeping oneself constantly informed of newly-available programs is a true challenge for the Bayesian teacher.

2. Rationality in economics and in statistics : the Bayesian connection

It is a remarkable fact that both micro-economic theory and decision theory, which provide a natural foundation for Bayesian methods, start by discussing the meaning of rational individual behaviour. Whereas elementary micro-economic theory considers decision-making in a riskless environment, decision theory is essentially a theory of decision-making under uncertainty. Thus, it is particularly appropriate that students in economics should have their foundation course in statistics presented to them as a natural extension of their introductory course in economics.

We sketch here only some of the ideas that we believe best illustrate the topic of this section and that are most likely to be fruitful as part of a pedagogical strategy for integrating the concept of rationality in economics and in statistics.

Uncertainty is usually introduced in intermediate or advanced courses on micro-economic theory where economic agents are typically assumed to process information in a Bayesian manner. This fact has several implications. Firstly, uncertainty is introduced simultaneously with the idea that agents (consumers or companies) either have different information initially or may receive it subsequently. This kind of economic model illustrates to the students why probabilities, representing information states, are personal or subjective, an idea they were first exposed to in their first course in (Bayesian) statistics.

In more advanced economics courses, the treatment of new information is endogenized and it is shown that the most efficient way of processing new information is to evaluate conditional probabilities - a further argument for using Bayes' theorem as a learning rule. Economic theory is also concerned with the rationality of group decision-making. Here again Bayesian ideas are useful in discussing the aggregation of preferences or opinions. Finally, we refer to recent developments in economics, such as the theory of rational expectations, the theory of contracts with asymmetric information or models relying on stochastic games which use the concept of perfect Bayesian equilibrium. All these developments provide opportunities to show that incorporating Bayesian ideas in the foundation statistics course does more than enrich students' conception of statistics: it provides an indispensable preparation for advanced economic analysis. This benefit is underlined in a particularly striking manner by recent developments in fields such as industrial organisation, where economic processes are studied using the formal tools of game theory, in which agents use information in a Bayesian way.

This topic has been the subject of an abundant literature and this report is not an appropriate outlet for reviewing that material. One might cite, for example, the fourth part of the volume edited by Zellner (1980) which collects interesting contributions by Grossman (1980), Cyert and De Groot (1980), and Prescott and Townsend (1980); see also Cyert and DeGroot (1987) or Gremaq (1988). The major international journals in economics and econometrics frequently publish contributions to this field and can provide a natural basis for establishing a reading list.

3. Observing economic reality

3.1 *Modelling economic phenomena*

A crucial role of statistical methodology is to offer guidelines when building models for the analysis of economic data. This target seems more relevant, but also more difficult, than providing inference rules for given models. Thus the topic of this section deserves, in our opinion, the most careful attention.

In econometric modelling, the role of economic theory is typically to provide constraints on standard statistical models. This is a well-known issue, for instance, in modelling cost functions (see, for example, Ilmakunnas, 1985) or demand systems, or in building simultaneous equations models understood as constrained multivariate regression models. "Testing" economic theory ultimately boils down to testing constraints.

The use of traditional testing procedures leads to well-known paradoxes: constraints are typically rejected once the sample size is large enough. Also, the fact that no model is deemed to be "perfect" (due to missing variables, unsatisfactory dynamic specification, etc.) further increases the probability of invalidly rejecting constraints. This often renders the interpretation of the significance level of an observation difficult, particularly when the choice of the null and alternative hypotheses is questionable or when a symmetric treatment of the two hypotheses seems to be more appropriate. Furthermore, a precise evaluation of procedures based on pre-testing is either extremely involved or not operational in practice.

Bayesian procedures then become particularly attractive for they show students that such pathologies are not necessarily intrinsic to empirical work. More specifically, Bayesian procedures allow one to relax constraints by imposing them "on the average" in a probabilistic framework, rather than "dogmatically" (i.e. with probability one). In particular, relaxing restrictions on nuisance parameters often avoids spoiling inference on parameters of interest when the alternative is either to introduce restrictions likely to be "not exactly true" or to dilute statistical information on too large a parameter space. Moreover, genuinely Bayesian testing procedures, such as those suggested in Florens and Mouchart (1989), may be more relevant as they are based on a comparison of the posterior distributions, with and without the constraint of quantities that may be functions of the parameter and of some observable variables (as is the case with elasticities which are functions of the cost, as in Ilmakunnas (1985)).

Many macro-econometric models are large models, involving many parameters, even though the sample sizes may be rather moderate. It eventually becomes of crucial importance to introduce substantial prior information if one wants to avoid trivial inferences (such as uselessly wide confidence intervals). In sizeable structural models, overidentifying restrictions may be required to avoid problems of overfitting *and* for computational reasons. But if we now consider VAR models, which are widely used in economic time series analysis, it may be observed that Litterman (1984), for instance, introduces prior distributions centred on random walks, thus a constraint in probabilistic form rather than with probability one. Such procedures are very much in use because of a two-fold advantage: the prior distribution is easily specified and understood and the posterior distribution is easily evaluated (in VAR models, RATS, for instance, does the job). These features enable students to think seriously about the flexibility of Bayesian procedures in facing genuine challenges of econometric modelling.

3.2 *Reporting statistical analysis*

Presenting the results of a statistical analysis in a Bayesian format appears to be easier for students in management or in economics to understand. Indeed, the concept of a posterior distribution is more easily, and correctly, understood than that of a sampling distribution. Thus, the Bayesian approach avoids the headache traditionally caused by trying to teach a proper interpretation of hypothesis testing or of confidence intervals in a sampling theory framework, not to mention the analysis of power in hypothesis testing.

This is so simply because, for practical purposes, the Bayesian format for presenting the results of a statistical analysis seems more natural and more relevant than that of sampling theory, particularly in situations where actual decisions must be taken. Furthermore, in applied general equilibrium models, for instance, redoing the analysis in a Bayesian way has proved to be highly illuminating and may reveal, by sheer experience, that what one was rejecting was not the model one was testing but rather the maintained hypothesis.

For more advanced students, discussions about reporting statistical analysis and eliciting prior information may provide a suitable opportunity to introduce recent ideas on the possibility of constructing inferences in the form of a family of posterior distributions. Such families are typically constructed from an (often non-parametric) family of prior distributions, thereby introducing some crucial issues with respect to robust Bayesian inferences.

4. Some difficulties

Systematic reliance on probabilistic arguments: One difficulty is of a conceptual nature. Bayesian statistics systematically relies on probabilistic arguments, the subtlety of which may cause difficulties to students in economics. More specifically, the central concepts of a conditional distribution and the conditional expectation are difficult to master both at an elementary level and at a more advanced level. Thus, in a first course, the idea that a conditional expectation is a random variable, or that a conditional distribution is also a random "object", typically creates problems. Similarly, it takes time to feel at ease with "marginal-conditional" decompositions. Note, however, that overcoming these difficulties promotes not only a good understanding of Bayesian methods, but also a better grasp of modern economic analysis.

At a less elementary level, more sophisticated tools are required when introducing more advanced econometric models. For instance, Dirichlet processes are naturally introduced when dealing with non-parametric methods, as are martingale methods in asymptotic theory, and using null sets correctly, as well as conditioning in infinite dimensional spaces, inevitably calls for some technicalities. Also, a reasonably deep analysis of the choice of the prior distribution, with emphasis on so-called "non-informative" prior specification, may require reference to concepts drawn from measure theory and invariance.

Probabilised parameters: Making parameters probabilistic does not, in our experience, raise too many difficulties as long as previous courses, following a sampling theory approach, have not been too narrowly conceived. Similarly, a subjective approach to the concept of probability, rather than a frequentist one, does not create particular problems for students in economics. This is very likely due to the fact that economics is naturally concerned with human behaviour, and so modern economic theory naturally accommodates such a subjectivistic approach.

General versus specific: We often have found it difficult to find a suitable compromise between two "evils". One extreme would be a purely conceptual course on general principles - as would flow, for instance, from narrowly following our monograph Florens, Mouchart and Rolin (1990). This monograph is oriented either to graduate (or advanced undergraduate) students in mathematical statistics or (pure or applied) mathematics, or to teachers in econometrics who wish to strengthen their knowledge of the foundations of Bayesian econometrics. Another extreme would be a course dealing with very specific models and involving catalogues of recipes describing how to manipulate cumbersome analytical expressions (such as, for instance, the conditional distributions of a matrix-student distribution). Regardless of how important such an operational culture may be, it seems to us that this can be better taught on an individual basis, as for instance with doctoral students, according to the actual needs of a specific research project.

A similar difficulty arises with respect to numerical problems. Regardless of how important these are when bringing a statistical analysis to a complete and operational conclusion, there is a serious danger of paying undue attention to numerical precision at the cost of impoverishing the statistical analysis.

Availability of computer programs: One of the major problems we encounter when teaching Bayesian methods is the difficulty of finding suitable computer programs. In the present world such programs are crucial if students are to be able to exploit their knowledge of Bayesian procedures and become aware of their genuine merits. Two types of difficulties are involved: one concerns the existence of such packages, and the other concerns their availability.

With regard to the first difficulty, although we are aware of several useful programs, we know of no program for Bayesian methods with a spectrum as wide as, for instance, SPSS, BMDP or SAS, or as user-friendly as, for instance, STATGRAPHICS, RATS or TSP, to mention just a minute sample of the packages presently available for carrying out sampling theory procedures. Note, however, that many Bayesian computations may be realised using these programs *provided* one limits oneself to non-informative or natural-conjugate prior distributions.

We have experimented fruitfully with programs developed at CORE (BRP and PTD). These programs are basically centered on the use of, and the numerical integration problems around, the poly-t distribution, a feature that makes these programs suitable only for rather advanced courses in Bayesian econometrics. These programs are presently being developed further by L Bauwens and M Lubrano of GREQE at the University of Aix-Marseille. They are seeking to develop a more user-friendly product which will accommodate both elementary and advanced procedures. The latest version, of which we have seen a demonstration, shows definite progress towards making these programs accessible to a much wider audience. The actual challenge of such programs is to incorporate powerful numerical integration procedures in as user-friendly a way as possible.

For the second difficulty, we have heard, often through informal channels, of many other Bayesian programs in the world. S J Press (1980, 1989), in particular, has made a very useful compilation, but our experience is that it is really difficult to keep oneself informed of the most important developments made around the world. Note, however, that such conferences as ICOTS, or the quadrennial Bayesian conferences in Valencia (Spain) are extremely useful for acquiring this kind of information.

All in all, we believe that the present situation is changing rapidly (in the desirable direction!) and that exchanging information, not only with respect to the existence of those programs but also concerning pedagogical experience when using them, is an important issue for the future.

Acknowledgement

Comments on earlier versions by V Barham, L Carter, D Cocchi, J H Drèze, E Moreno, J -M Rolin, E Sowey, and M Steel led us to deepen the treatment and to improve the presentation and, without implication, are gratefully acknowledged.

Appendix

It may be useful to give two quite different examples of course contents drawn from the authors' experience.

Example 1

Audience
5th university year ("DEA" : Diplôme d'Etudes Approfondies), for students in economics with a strong quantitative background.

Content
1. Bayesian models
   - definition and simple example of a probability on the product space "parameter × observation"
   - Bayesian statistics and decision theory
   - prior specification (natural conjugate and non-informative)

- reductions (marginalisation and conditioning, dually on the parameters and the observations)
- integration and numerical problems

2. Linear models
   - review of some probability distributions
   - univariate regression model with natural conjugate prior specification, multivariate extension
   - an example of limited information
   - extensions of the linear model (heteroscedasticity etc.)
3. Statistical procedures
   - statistical properties of posterior expectations
   - least squares approximations
   - Bayesian tests
4. Some general models
   - non-parametric models (Dirichlet process)
   - semi-parametric models
   - exponential models

Remark
In this course, sections on general principles (1 and 3) alternate with sections on more specific models (2 and 4).

Example 2

Audience
2nd university year, for students in economics or management with a rather quantitative orientation.

Content

1. Decisions and statistics
   - problems of decisions under uncertainty (Moral Expectation Theorem)
   - sequential decisions (extensive form and normal form analysis)
   - from decision analysis to statistical analysis (statistical modelling, decisional formalisation of statistical problems)
2. Point estimation
   - the problem of point estimation
   - Bayesian analysis (general case, quadratic loss, normal sampling)
   - statistical information (sufficiency, ancillarity, exponential family, Score and Fisher information)
   - prior specification (natural-conjugate, non-informative, others)
   - sampling theory analysis: general principles (small and large samples)
   - construction of point estimators (maximum likelihood, moments, others)
   - unbiased estimation (Rao-Cramer, Rao-Blackwell, Lehmann-Scheffé)
   - invariant estimation (in general, location and scale parameter)

3. Tests of hypotheses
   - the general problem of hypothesis testing
   - Bayesian analysis of two-decisions problems
   - sampling theory analysis: general principles
   - construction of testing procedures
   - unbiased tests
   - invariant tests
4. Analysis of variance
   - statistical model (fixed or random effect, 1 or 2 ways)
   - Bayesian analysis
   - sampling theory analysis

Remarks

In this course, a decisional framework is used to present both Bayesian and sampling theory approaches in a presentation that is intended to be as symmetric as possible. Experience suggests that some students do appreciate such a systematic approach, other students find the conceptual burden somewhat heavy.

References

Cyert, R M and De Groot, M H (1980) Learning applied to utility functions. In: *Bayesian Analysis in Econometrics and Statistics*, see Zellner (1980), 159-168.

Cyert, R M and De Groot, M H (1987) *Bayesian Analysis and Uncertainty in Economic Theory*. Rowman and Littlefield, Totowa (New Jersey).

Florens, J -P and Mouchart, M (1989) Bayesian specification tests. In: B Cornet and H Tulkens (eds) *Contributions in Operations Research and Economics*, 467-490. MIT Press, Cambridge.

Florens, J -P, Mouchart, M and Rolin, J -M (1990) *Elements of Bayesian Statistics*. Marcel Dekker, New York.

Gremaq, A A (1988) *Dynamique, Information Incomplète et Stratégies Industrielles*. Economica, Paris.

Grossman, S J (1980) Rational expectations and the econometric modelling of markets subject to uncertainty : a Bayesian approach. In: *Bayesian Analysis in Econometrics and Statistics*, see Zellner (1980), 143-158.

Ilmakunnas, P (1985) Bayesian estimation of cost functions with stochastic or exact constraints on parameters. *International Economic Review* **26**, 111-134.

Litterman, R B (1984) Forecasting and policy analysis with Bayesian autoregressive models. *Federal Reserve Bank of Minneapolis Quarterly Review* **Fall**.

Prescott, E C and Townsend, R M (1980) Equilibrium under uncertainty : multiagent statistical decision theory. In: *Bayesian Analysis in Econometrics and Statistics*, see Zellner (1980), 169-194.

Press, S J (1980) Bayesian computer programs. In: *Bayesian Analysis in Econometrics and Statistics*, see Zellner (1980), 429-442.

Press, S J (1989) *Bayesian Statistics : Principles, Models and Applications*. John Wiley, New York.

Zellner, A (ed) (1980) *Bayesian Analysis in Econometrics and Statistics. Essays in Honour of Harold Jeffreys*. North Holland, Amsterdam.