BAYESIAN STATISTICS

José M. Bernardo

Faculty of Mathematics, University of Valencia, Valencia, Spain

Keywords: Amount of Information, Axiomatics, Bayes Estimator, Bayes Factor, Bayes Theorem, Credible Region, Degree of Belief, Decision Theory, Exchangeability, Foundations of Statistics, Hierarchical Models, Hypothesis Testing, Interval Estimation, Intrinsic Discrepancy, Likelihood Principle, Lindley's Paradox, Logarithmic Divergence, Maximum Entropy, Model Choice, Model Criticism, Noninformative Priors, Nuisance Parameters, Objectivity, Point Estimation, Posterior Distribution, Predictive Distribution, Prior Choice, Prior Distribution, Probability Assessment, Probability model, Probability Theory, Reference Distributions, Representation Theorem, Scientific Reporting, Sensitivity Analysis, Steins's Paradox, Sufficiency.

Contents

Summary

Statistics is the study of uncertainty. The field of statistics is based on two major paradigms: conventional and Bayesian. Bayesian methods may be derived from an

axiomatic system and provide a *complete* paradigm for both statistical inference and decision making under uncertainty. Bayesian methods provide a *coherent* methodology which makes it possible to incorporate relevant initial information, and which alleviates many of the difficulties faced by conventional statistical methods. The Bayesian paradigm is based on an interpretation of probability as a *conditional measure of uncertainty* which closely matches the sense of the word 'probability' in ordinary language. Statistical inference about a quantity of interest is described as a modification of the uncertainty about its value in the light of evidence, and Bayes theorem specifies how this modification should be made. Bayesian methods may be applied to highly structured complex problems, which have been often not easily tractable by traditional statistical methods. The special situation, often met in scientific reporting and public decision making, where the only acceptable information is that which may be deduced from available documented data, is addressed as an important particular case.

1. Introduction

Scientific, experimental or observational results generally consist of (possibly many) sets of data of the general form $D = \{\boldsymbol{x}_1, \ldots.., \boldsymbol{x}_n\}$, where the $\boldsymbol{x}_i$s are somewhat "homogeneous" (possibly multidimensional) observations $\boldsymbol{x}_i$. Statistical methods are then typically used to derive conclusions on both the nature of the process which has produced those observations, and on the expected behavior at future instances of the same process. A central element of *any* statistical analysis is the specification of a *probability model* which is assumed to describe the mechanism which has generated the observed data D as a function of a (possibly multidimensional) parameter (vector) $\boldsymbol{\omega} \in \Omega$, sometimes referred to as the *state of nature*, about whose value only limited information (if any) is available. All derived statistical conclusions are obviously conditional on the assumed probability model.

Unlike most other branches of mathematics, conventional methods of statistical inference suffer from the lack of an axiomatic basis; as a consequence, their proposed desiderata are often mutually incompatible, and the analysis of the same data may well lead to incompatible results when different, apparently intuitive procedures are tried (see the 1970's monographs by Lindley and by Jaynes for many instructive examples). In marked contrast, the Bayesian approach to statistical inference is firmly based on axiomatic foundations which provide a unifying logical structure, and guarantee the mutual consistency of the methods proposed. Bayesian methods constitute a *complete* paradigm to statistical inference, a scientific revolution in Kuhn's sense.

Bayesian statistics only require the *mathematics* of probability theory and the *interpretation* of probability which most closely corresponds to the standard use of this word in everyday language: it is no accident that some of the more important seminal books on Bayesian statistics, such as the works of de Laplace, de Finetti or Jeffreys, are actually entitled "Probability Theory". The practical consequences of adopting the Bayesian paradigm are far reaching. Indeed, Bayesian methods (i) reduce statistical inference to problems in probability theory, thereby minimizing the need for completely new concepts, and (ii) serve to discriminate among conventional statistical techniques, by either providing a logical justification to some (and making explicit the conditions under which they are valid), or proving the logical inconsistency of others.

The main consequence of these foundations is the mathematical *need* to describe by means of probability distributions all uncertainties present in the problem. In particular, unknown parameters in probability models *must* have a joint probability distribution which describes the available information about their values; this is often regarded as *the* characteristic element of a Bayesian approach. Notice that (in sharp contrast to conventional statistics) *parameters are treated as random variables* within the Bayesian paradigm. This is not a description of their variability (parameters are typically *fixed unknown* quantities) but a description of the uncertainty about their true values.

An important particular case arises when either no relevant prior information is readily available, or that information is subjective and an "objective" analysis is desired, one that is exclusively based on accepted model assumptions and well-documented data. This is addressed by *reference analysis* which uses information-theoretic concepts to derive appropriate reference posterior distributions, defined to encapsulate inferential conclusions on the quantities of interest, solely based on the supposed model and the observed data.

In this article it is assumed that probability distributions may be described through their probability density functions, and no distinction is made between a random quantity and the particular values that it may take. Bold italic roman fonts are used for *observable* random vectors (typically data) and bold italic Greek fonts are used for unobservable random vectors (typically parameters); lower case is used for variables and upper case for their dominion sets. Moreover, the standard mathematical convention of referring to *functions*, say f and g of $\boldsymbol{x} \in X$, respectively by $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, will be used throughout. Thus, $p(\boldsymbol{\theta} \mid C)$ and $p(\boldsymbol{x} \mid C)$ respectively represent general *probability densities* of the random vectors $\boldsymbol{\theta} \in \Theta$ and $\boldsymbol{x} \in X$ under conditions *C*, so that $p(\boldsymbol{\theta} \mid C) \geq 0, \int_\Theta p(\boldsymbol{\theta} \mid C) d\boldsymbol{\theta} = 1,$ and $p(\boldsymbol{x} \mid C) \geq 0, \int_X p(\boldsymbol{x} \mid C) d\boldsymbol{x} = 1.$ This admittedly imprecise notation will greatly simplify the exposition. If the random vectors are discrete, these functions naturally become probability mass functions, and integrals over their values become sums.

Density functions of specific distributions are denoted by appropriate names. Thus, if x is a random quantity with a normal distribution of mean μ and standard deviation σ, its probability density function will be denoted $\mathrm{N}(x \mid \mu, \sigma)$. Table 1 contains definitions of other distributions used in this article.

Name	*Probability Density or Probability Mass Function*	*Parameter(s)*
Beta	$\mathrm{Be}(x \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, x \in (0,1)$	$\alpha > 0, \beta > 0$
Binomial	$\mathrm{Bi}(x \mid n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, x \in \{0, \ldots\ldots, n\}$	$n \in \{1, 2, \ldots.\}, \theta \in (0,1)$
Exponential	$\mathrm{Ex}(x \mid \theta) = \theta e^{-\theta x}, x > 0$	$\theta > 0$

ExpGamma	$\mathrm{Eg}(x\|\alpha,\beta)=\frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}},x>0$	$\alpha>0,\beta>0$
Gamma	$\mathrm{Ga}(x\|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x},x>0$	$\alpha>0,\beta>0$
NegBinomial	$\mathrm{Nb}(x\|r,\theta)=\theta^{r}\binom{r+x-1}{r-1}(1-\theta)^{x},x\in\{0,1,\ldots.\}$	$r\in\{1,2,\ldots.\},\theta\in(0,1)$
Normal	$\mathrm{N}_k(\boldsymbol{x}\|\boldsymbol{\mu},\Sigma)=\frac{\|\Sigma\|^{-1/2}}{(2\pi)^{k/2}}\exp[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{t}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})],\boldsymbol{x}\in\Re^{k}$	$\boldsymbol{\mu}\in\Re^{k},\Sigma$ def.pos.
Poisson	$\mathrm{Pn}(x\|\lambda)=e^{-\lambda}\frac{\lambda^{x}}{x!},x\in\{0,1,\ldots..\}$	$\lambda>0$
Student	$\mathrm{St}(x\|\mu,\sigma,\alpha)=\frac{\Gamma(\frac{\alpha+1}{2})}{\Gamma(\frac{\alpha}{2})}\frac{1}{\sigma\sqrt{\alpha\pi}}[1+\frac{1}{\alpha}(\frac{x-\mu}{\sigma})^{2}]^{-(\alpha+1)/2},x\in\Re$	$\mu\in\Re,\sigma>0,\alpha>0$

Table 1. Notation for common probability density and probability mass functions

Bayesian methods make frequent use of the concept of logarithmic divergence, a very general measure of the goodness of the approximation of a probability density $p(\boldsymbol{x})$ by another density $\hat{p}(\boldsymbol{x})$. The *logarithmic divergence* of a probability density $\hat{p}(\boldsymbol{x})$ of the random vector $\boldsymbol{x}\in X$ from its true probability $p(\boldsymbol{x})$, is defined as $\delta\{\hat{p}(\boldsymbol{x})\,|\,p(\boldsymbol{x})\}=\int_X p(\boldsymbol{x})\log\{p(\boldsymbol{x})/\hat{p}(\boldsymbol{x})\}d\boldsymbol{x}$. It may be shown that (i) the logarithmic divergence is non-negative (and it is zero, if, and only if, $\hat{p}(\boldsymbol{x})=p(\boldsymbol{x})$ almost everywhere), and (ii) that $\delta\{\hat{p}(\boldsymbol{x})\,|\,p(\boldsymbol{x})\}$ is invariant under one-to-one transformations of $\boldsymbol{x}$.

This article contains a brief summary of the mathematical foundations of Bayesian statistical methods (Section 2), an overview of the paradigm (Section 3), a description of useful inference summaries, including estimation and hypothesis testing (Section 4), an explicit discussion of objective Bayesian methods (Section 5), a detailed analysis of a simplified case study (Section 6), and a final discussion which includes pointers to further issues not addressed here (Section 7).

2. Foundations

A central element of the Bayesian paradigm is the use of probability distributions to describe all relevant unknown quantities, interpreting the probability of an event as a conditional measure of uncertainty, on a [0,1] scale, about the occurrence of the event in some specific conditions. The limiting extreme values 0 and 1, which are typically inaccessible in applications, respectively describe impossibility and certainty of the occurrence of the event.

This interpretation of probability includes and extends all other probability interpretations. There are two independent arguments which prove the mathematical inevitability of the use of probability distributions to describe uncertainties; these are summarized later in this section.

2.1. Probability as a Measure of Conditional Uncertainty

Bayesian statistics uses the word *probability* in the same sense in which this word is used in everyday language, as a *conditional measure of uncertainty* associated with the occurrence of particular event, given the available information and the accepted assumptions. Thus, $\Pr(E|C)$ is a measure of (presumably rational) belief in the occurrence of the *event E* under *conditions C*. It is important to stress that probability is *always* a function of two arguments, the event *E* whose uncertainty is being measured, and the conditions *C* under which the measurement takes place; "absolute" probabilities do not exist. In typical applications, one is interested in the probability of some event *E* given the available *data D*, the set of *assumptions A* which one is prepared to make about the mechanism which has generated the data, and the relevant contextual *knowledge K*, which might be available. Thus, $\Pr(E|D,A,K)$ is to be interpreted as a measure of (presumably rational) belief in the occurrence of the *event E*, given data *D*, assumptions *A* and any other available knowledge *K*, as a measure of how "likely" is the occurrence of *E* in these conditions. Sometimes, but certainly not always, the probability of an event under given conditions may be associated with the relative frequency of "similar" events in "similar" conditions. The following examples are intended to illustrate the use of probability as a conditional measure of uncertainty.

Probabilistic diagnosis. A human population is known to contain 0.2% of people infected by a particular virus. A person, *randomly selected* from that population, is subject to a test which is from laboratory data known to yield positive results in 98% of infected people and in 1% of non- infected, so that, if *V* denotes the event that a person carries the virus and + denotes a positive result, $\Pr(+|V)$=0.98 and $\Pr(+|\overline{V})$=0.01. Suppose that the result of the test turns out to be positive. Clearly, one is then interested in $\Pr(V|+,A,K)$, the *probability* that the person carries the virus, given the positive result, the assumptions *A* about the probability mechanism generating the test results, and the available knowledge *K* of the prevalence of the infection in the population under study (described here by $\Pr(V|K)=0.002$) . An elementary exercise in probability algebra, which involves Bayes theorem in its simplest form (see Section 3), yields $\Pr(V|+,A,K)=0.164$. Notice that the four probabilities involved in the problem have *the same interpretation*: they are all conditional measures of uncertainty. Besides, $\Pr(V|+,A,K)$ is *both* a measure of the uncertainty associated with the event that the particular person who tested positive is actually infected, *and* an *estimate* of the proportion of people in that population (about 16.4%) that would eventually prove to be infected among those which yielded a positive test.

Estimation of a proportion. A survey is conducted to estimate the proportion θ of individuals in a population who share a given property. A random sample of n elements is analyzed, *r* of which are found to possess that property. One is then typically interested in using the results from the sample to establish regions of [0,1] where the unknown value of θ may plausibly be expected to lie; this information is provided by *probabilities* of the form $\Pr(a<\theta,b|r,n,A,K)$, a conditional measure of the uncertainty about the event that θ belongs to (a,b) *given* the information provided by the data (r,n), the assumptions *A* made on the behavior of the mechanism which has generated the data (a random sample of n Bernoulli trials), and any relevant knowledge *K* on the values of θ which might be available. For example, after a political survey in which 720 citizens out of a random sample of 1500 have declared their support to a particular political measure, one may conclude that $\Pr(\theta < 0.5 \mid 720, 1500, A, K) = 0.933$, indicating a probability of about 93% that a referendum of that issue would be lost. Similarly, after a screening test for an infection where 100 people have been tested, none of which has turned out to be infected, one may conclude that $\Pr(\theta < 0.01 \mid 0, 100, A, K) = 0.844$, or a probability of about 84% that the proportion of infected people is smaller than 1%.

Measurement of a physical constant. A team of scientists, intending to establish the unknown value of a physical constant μ, obtain data $D = \{x_1, \ldots, x_n\}$ which are considered to be measurements of μ subject to error. The probabilities of interest are then typically of the form $\Pr(a < \mu < b \mid x_1, \ldots, x_n, A, K)$, the *probability* that the unknown value of μ (fixed in nature, but unknown to the scientists) lies within an interval (a,b) given the information provided by the data *D*, the assumptions *A* made on the behavior of the measurements mechanism, and whatever knowledge *K* might be available on the value of the constant μ. Again, those probabilities are conditional measures of uncertainty which describe the (necessarily probabilistic) conclusions of the scientists on the true value of μ, given available information and accepted assumptions. For example, after a classroom experiment to measure the gravitational field with a pendulum, a student may report (in m/sec^2) something like $\Pr(9.788 < g < 9.829 \mid D, A, K) = 0.95$, meaning that, under accepted knowledge *K* and assumptions *A*, the *observed* data *D* indicate that the true value of g lies within 9.788 and 9.829 with probability 0.95, a conditional uncertainty measure on a [0,1] scale. This is naturally compatible with the fact that the value of the gravitational field at the laboratory may well be known with high precision from available literature or from precise previous experiments, but the student may have been instructed *not* to use that information as part of the accepted knowledge *K*, Under some conditions, it is also true that if the same *procedure* were actually used by many other students with similarly obtained data sets, their reported intervals would actually cover the true value of g in approximately 95% of the cases, thus providing some form of *calibration* for the student's probability statement (see Section 5.2).

Prediction. An experiment is made to count the number *r* of times that an event *E* takes place in each of n replications of a well defined situation; it is observed that *E* does take place r_i times in replication *i*, and it is desired to forecast the number of times *r* that *E* will take place in a future, similar situation. This is a *prediction* problem on the value of an *observable* (discrete) quantity *r*, given the information provided by data *D*, accepted assumptions *A* on the probability mechanism which generate the r_i's , and any relevant available knowledge *K*. Hence, simply the computation of the probabilities $\{\Pr(r|r_1,....,r_n,A,K)\}$, for r = 0,1,…., is required. For example, the quality assurance engineer of a firm which produces automobile restraint systems may report something like $\Pr(r=0|r_1=....=r_{10}=0,A,K)=0.953$ after observing that the entire production of airbags in each of n =10 consecutive months has yielded no complaints from their clients.

This should be regarded as a measure, on a [0,1] scale, of the conditional uncertainty, given observed data, accepted assumptions and contextual knowledge, associated with the event that no airbag complaint will come from next month's production and, if conditions remain constant, this is also an estimate of the proportion of months expected to share this desirable property.

A similar problem may naturally be posed with continuous observables. For instance, after measuring some continuous magnitude in each of n randomly chosen elements within a population it may be desired to forecast the proportion of items in the whole population whose magnitude satisfies some precise specifications.

As an example, after measuring the breaking strengths $\{x_1,.....,x_{10}\}$ of 10 randomly chosen safety belt webbings to verify whether or not they satisfy the requirements of remaining above 26 kN, the quality assurance engineer may report something like $\Pr(x>26|x_1,....,x_{10},A,K)=0.9987$.

This should be regarded as a measure, on a [0,1] scale, of the conditional uncertainty (given observed data, accepted assumptions and contextual knowledge) associated with the event that a randomly chosen safety belt webbing will support no less than 26 kN. If production conditions remain constant, it will also be an estimate of the proportion of safety belts which will conform to this particular specification.

Often, additional information of future observations is provided by related covariates. For instance, after observing the outputs $\{\boldsymbol{y}_1,.....,\boldsymbol{y}_n\}$ which correspond to a sequence $\{\boldsymbol{x}_1,.....,\boldsymbol{x}_n\}$ of different productions conditions, it may be desired to forecast the output $\boldsymbol{y}$ which would correspond to a particular set $\boldsymbol{x}$ of productions conditions.

For instance, the viscosity of commercial condensed milk is required to be within specified values *a* and *b*; after measuring the viscosities $\{y_1.....,y_n\}$ which correspond to samples of condensed milk produced under different physical conditions $\{\boldsymbol{x}_1,.....,\boldsymbol{x}_n\}$, production engineers will require probabilities of the form $\Pr(a<y<b|\boldsymbol{x},(y_1,\boldsymbol{x}_1),....,(y_n,\boldsymbol{x}_n),A,K)$.

This is a conditional measure of the uncertainty (always given observed data, accepted assumptions and contextual knowledge) associated with the event that condensed milk produced under conditions $\boldsymbol{x}$ will actually satisfy the required viscosity specifications.

-
-
-

Bibliography

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Berlin: Springer. [A thorough account of Bayesian methods emphasizing its decision-theoretical aspects]

Bernardo, J.M. (1979). Expected information as expected utility. *Ann*. *Statist*. **7**, 686-690. [Establishes statistical inferences as a decision problem with an information-based utility function]

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *J. R. Statist. Soc. B* **41**, 113-147 (with discussion). Reprinted in Bayesian Inference 1 (G. C. Tiao and N. G. Polson, eds). Oxford: Edward Elgar, 229-263.[The original paper on reference analysis]

Bernardo, J.M. (1997). Noninformative priors do not exist. *J. Statist. Planning and Inference* **65**, 159-189 (with discussion). [A non-technical analysis of the polemic on objective Bayesian statistics]

Bernardo, J. M. and Rueda, R. (2002). Bayesian hypothesis testing: A reference approach. *International Statistical Review* **70** , 351-372. [A new decision-oriented approach to sharp hypothesis testing]

Bernardo, J.O.Berger, A.P. Dawid and A.F.M. Smith, eds). Oxford: University Press, 101-130 (with discussion). [A decision-oriented approach to sharp hypothesis testing]

Bernardo, J.M. and Ramón, J.M. (1998). An introduction to Bayesian reference analysis: inference on the ratio of multinomial parameters. *The Statistician* **47**, 1-35. [An elementary introduction to objective Bayesian analysis]

Bernardo, J.M. and Smith, A.F.M. (1994) *Bayesian Theory*, Chichester: Wiley. [A thorough account of key concepts and theoretical results in Bayesian statistics at a graduate level, with extensive bibliography]

Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P. Heckerman, D., Smith, A.F.M. and West, M. (Eds.) (2003). *Bayesian Statistics 7*. Oxford: Oxford University Press. [The Proceedings of the 7th Valencia International Meeting on Bayesian Statistics; the Valencia meetings, held every four years, provide definite up-do-date overviews on current research within the Bayesian paradigm.]

Bernardo, J. M. and Rueda, R. (2005). Intrinsic credible regions: An objective Bayesian Approach to interval estimation. Test 14,317-384 [A modern approach to objective Bayesian region estimation, with discussion from many eminent statisticians]

Berry, D.A. (1996). *Statistics, a Bayesian Perspective*, Belmont, CA: Wadsworth. [A very good introduction to Bayesian statistics from a subjectivist viewpoint]

Box, G.E.P. and Tiao, G.C. (1973) . *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley. [An excellent objective Bayesian account of standard statistical problems]

Dawid, A. P., Stone, M. and Zidek, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference. *J. R. Statist. Soc. B* **35,** 189-233 (with discussion). [Proves that a unique “non-informative” prior for all parameters of interest within a given model is not possible]

DeGroot, M.H. (1970). *Optimal Statistical Decisions*, New York: McGraw-Hill. [A thorough account of Bayesian decision theory and Bayesian inference with a rigorous treatment of foundations]

Efron, B. (1986). Why isn’t everyone a Bayesian? *Amer*. *Statist*. **40,** 1-11 (with discussion). [A good example of the polemic between Bayesian and non-Bayesian approaches to statistics] de. Finetti, B. (1970). *Teoria delle Probabilitá*, Turin: Einaudi. English translation as "*Theory of Probability* in 1975, Chichester: Wiley. [An outstanding book on probability and statistics from a subjective viewpoint]

Geisser,S. (1993). *Predictive Inference: an Introduction*. London: Chapman and Hall. [A comparative account of frequentist and objective Bayesian methods of prediction]

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. *J.Amer. Statist. Assoc.* **85,** 398-409. [An excellent primer on simulation-based techniques to numerical integration in the context of Bayesian statistics]

Gelman, A., Carlin, J.B., Stern, H. and Rubin, D.B. (1995). *Bayesian Data Analysis*. London: Chapman and Hall. [A comprehensive treatment of Bayesian data analysis emphasizing computational tools]

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall. [An excellent introduction to MCMC methods and their applications]

Jaynes, E. T. (1976). Confidence intervals vs. Bayesian intervals. *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science 2* (W.L. Harper and and C.A. Hooker, eds). Dordrecht: Reidel, 175--257 (with discussion)

Kass, R.E. and Raftery, A.E. (1995) . Bayes factors, *J.Amer. Statist. Assoc. 90*, 773-795. [A very good review of Bayes factor methods for hypothesis testing]

Lindley, D.V. (1972). *Bayesian Statistics, a Review*. Philadelphia, PA: SIAM. [A sharp comprehensive review of the whole subject up to the 1970s', emphasizing its internal consistency]

Lindley, D.V. (1990). The 1988 Wald memorial lecture: The present position in Bayesian Statistics. *Statist. Sci.* **5**, 44-89 (with discussion). [An informative account of the Bayesian paradigm and its relationship with other attitudes to inference]

Lindley, D.V. (2000). The philosophy of Statistics. *The Statistician* **49**, 293-337 (with discussion). [A recent description of the Bayesian paradigm from a subjectivist viewpoint]

O'Hagan, A. (1994). *Bayesian Inference* London: Edward Arnold. [A good account of Bayesian inference integrated into Kendall's Library of Statistics]

Press, S.J. (1972). *Applied Multivariate Analysis: using Bayesian and Frequentist Methods of Inference*. Melbourne, Fl: Krieger. [A comprehensive and comparative account of frequentist and objective Bayesian methods of inference in multivariate problems]

West, M. and Harrison, P.J. (1989). *Bayesian Forecasting and Dynamic Models*. Berlin: Springer. [An excellent thorough account of Bayesian time series analysis]

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley. Reprinted in 1987, Melbourne, FL: Krieger. [A detailed objective Bayesian analysis of linear models]

Biographical Sketch

José-Miguel Bernardo was born in Valencia (Spain) on March 12th, 1950. In 1974 he got a Ph.D. in Mathematics at University of Valencia; in 1976 he got a second Ph.D. in Statistics at University College London. Since 1978 he is Professor of Statistics (Catedrático) at the University of Valencia, Spain. Between 1989 and 1993, he was part of the State Government of Valencia as General Director of Decision Analysis. He is Fellow of the American Statistical Association, member of the Real Academia de Ciencias de Madrid, and member of the International Statistical Institute. He was Founding Co-President of the International Society for Bayesian Analysis. He was Founding Editor of Test and is, or has been, Associate Editor of Estadística Española, Journal of the Royal Statistical Society (Series B), The Statistician, Current Index of Statistics, and Statistical Theory and Methods Abstracts. He regularly referees projects for the Spanish Scientific and Technological Commission (Madrid) and for the National Science Foundation (Washington). He is Organizer and Programme Committee Member of the Valencia International Meetings on Bayesian Statistics (held in 1979, 1983, 1987, 1991, 1994, 1998, 2002 and 2006), established world forums on Bayesian Methods, sponsored every four years by the University of Valencia. He is the author of Bioestadística, una Perspectiva Bayesiana, Barcelona: Vicens-Vives, 1981 and coauthor (with A. F. M. Smith) of Bayesian Theory, (Chichester: Wiley, 2007 (2^{nd} edition)). He has been co-editor of the Proceedings of the Valencia International Meetings on Bayesian Statistics, currently published by Oxford University Press. He has published over 70 research papers.