Scientific Reasoning: The Bayesian Approach
Third Edition, 2006
By Colin Howson and Peter Urbach

Review by Glenn Shafer, Rutgers University and Royal Holloway University of London
July 27, 2006

The first edition of this book appeared in 1989, the second in 1993. The appearance of a third edition is testimony to the book's value to a large number of potential readers. Its authors are philosophers at LSE, and it seems likely that its readers, too, have mostly been philosophers. It remains the best presentation for philosophers of the Bayesian viewpoint on statistical inference. Although it proceeds in a relatively verbal way, it gets as far as the bivariate normal distribution in its review of probability theory, and its treatment of statistical inference covers significance testing, confidence intervals, randomization, clinical trials, regression, prediction, data analysis, and stable estimation.

Statisticians, even those already familiar with the arguments of 20th century Bayesians such as Jeffreys, Good, Savage, Lindley, and de Finetti, may also find the book informative. One of its nine chapters, Chapter 4, is devoted to a Bayesian treatment of issues that are not commonly discussed in the statistical literature because they concern deterministic rather than stochastic models. Even when discussing more familiar issues, the book cites contributions by philosophers not commonly read by statisticians (David Lewis, for example), and in some cases the authors propound original views.

The authors are Bayesians of a practical sort. They readily acknowledge the existence of two distinct kinds of probabilities – objective and epistemic. They take von Mises (and Church; they make no reference to Wald or Ville) as their authority on objective probabilities. Epistemic probabilities, they explain, are "personal fair betting quotients" – not rates at which you would really bet, for this might involve utilities, but rates at which you think it fair to bet. They do not envision a Bayesian agent who has probabilities for everything and conditions these probabilities as information comes in. Their focus is more on Bayes's Theorem, which they see as useful for assessing evidence *e* even when it is not the latest thing you have learned. Even if you knew *e* first, it may make sense, they think, to assess probabilities based on your other evidence and then take *e* into account by Bayes's Theorem. The goal of the book seems to be to demonstrate by example that this is invariably the most enlightening way to proceed.

They do not believe in "ignorance priors" as Harold Jeffreys did. But they regard the probability axioms as rules for reasoning, analogous to the rules of logic. In previous editions they included countable additivity as one of these rules of reasoning, but now they have decided against this. As one who has long been interested in assessing evidence using degrees of belief that are not even finitely additive, I must sadly note that the authors do not deign to mention the literature on this topic.

Other statisticians will complain that the authors often stop short of mentioning relatively technical but insightful aspects of the issues they discuss. For example, they criticize the Gauss-Markov theorem by evoking the possibility of trading bias for variance, but they do not go on to mention insights provided by authors such as Wald and Stein. They make a formal argument against confidence coefficients being interpreted as probabilities, but they do not explain what makes these formal arguments persuasive for statisticians – the existence of relevant subsets in Fisher's sense.

A greater shortcoming is that the third edition does not really bring the picture of Bayesian statistics up-to-date. In their preface, the authors point out how sharply the use of Bayesian methods shot up in the 1990s. But they do not explore the context of this paradigm shift and the new philosophical issues it raises. Bayesian methods have become popular because we now have the computational means to implement them in very complex models where classical Neyman-Pearson methods are hopeless. But as the models grow in complexity and numbers of parameters, the meaningfulness of prior distributions for these parameters becomes increasingly elusive, and the alternatives are now new methods in machine learning, not the classical methods the authors criticize.

For my part, I also fault the book for a failing that it shares with most contemporary literature on the foundations of probability and statistics – its unspoken but relentless assumption that only what appeared in English was important in the development of the ideas of probability and statistical inference. Bayes is mentioned more often than Laplace. Pearson and Fisher seem to materialize from nowhere. The names Borel, Fréchet, Lévy, and Ville never appear. Cournot's principle is mentioned and decried, but with no acknowledgement of the subtlety of the giants who defended it. The low point came for me when the authors cited Kolmogorov's "1950 monograph." The reader can learn from the bibliography that this monograph was published in 1933 in German. But it is evidently only the appearance of the English translation in 1950 that counts.

So be it. The book is not history, and it does not tell us about the present or the future. But it is a readable introduction, especially for those philosophically inclined, to what was going on in English-language debates about the foundations of probability and statistics in the mid-20th century.