Statistical Science
2009, Vol. 24, No. 2, 173–175
DOI: 10.1214/09-STS284E
Main article DOI: 10.1214/09-STS284

Comment

José M. Bernardo

The authors provide an authoritative lecture guide of *Theory of Probability*, where they clearly state that the more useful material today is that contained in Chapters 3 and 5, which respectively deal with *estimation*, and *hypothesis testing*. We argue that, from a contemporary viewpoint, the impact of Jeffreys proposals on those two problems is rather different, and we describe what we perceive to be the state of the question nowadays, suggesting that Jeffreys's dramatically different treatment is not necessary, and that a joint objective approach to those two problems is indeed possible.

1. INTRODUCTION

As the authors point out, *Theory of Probability* is an indispensable, if often difficult to navigate, Bayesian foundational text. Their authoritative lecture guide is therefore very welcome. As should be clear from their review, the main useful material today is contained in Chapters 3 and 5 which, respectively, deal with *estimation*, in the sense of deriving an objective posterior distribution for the quantity of interest, and *hypothesis testing*, presented as a derivation of an objective posterior probability for the hypothesis under consideration. I believe that, from a contemporary viewpoint, the impact of Jeffreys proposals on those two problems is rather different, as I now briefly try to describe.

2. ESTIMATION

One-parameter Jeffreys estimation prior (Jeffreys rule). Following his own pioneering work (Jeffreys, 1946), the book introduces in Section 3.10 what it is now considered the main meaning of the confusing denomination "Jeffreys prior." Thus, to obtain an objective posterior density for the parameter α of a probability model $f(x|\alpha)$, he proposes the formal use in Bayes theorem of the (often improper) prior $\pi(\alpha) \propto |I(\alpha)|^{1/2}$, where $I(\alpha)$ is Fisher information function. As the authors point out, Jeffreys's motivation is rather obscure: he describes $I(\alpha)$ as a second order approximation to two functional distances, and notes that $|I(\alpha)|^{1/2}$ happens to be invariant under one-to-one transformations. No trace of its more intuitive interpretation in terms of the prior which assigns equal probabilities to equally distinguishable subregions of the parameter space (Lindley, 1961). Also, even in its third (1961) edition, the book only gives a cursory reference to the independent, essentially simultaneous, derivation of the same "rule" produced by Perks (1947) in a much underrated paper. That said, Jeffreys (or Jeffreys–Perks) rule is today the objective prior of choice for regular problems with one continuous parameter, and has been justified in this simple case from many different viewpoints, including coverage properties (Welch and Peers, 1963), minimum bias (Hartigan, 1965), data translation (Box and Tiao, 1973) and information-theoretical arguments (Bernardo, 1979; Berger, Bernardo and Sun, 2009). In one-parameter problems, Jeffreys left without solution non-regular models (e.g., those where the sampling space depends on the parameter) and models with a discrete parameter (although he suggested a very interesting hierarchical argument to deal with the particular example of the hypergeometric distribution).

Many-parameter Jeffreys estimation prior (multiparameter Jeffreys rule). The arguments used to propose his rule for one continuous parameter regular models extend to the corresponding multiparameter case, leading to $\pi(\boldsymbol{\alpha}) \propto |I(\boldsymbol{\alpha})|^{1/2}$, where $I(\boldsymbol{\alpha})$ is now Fisher information matrix. As the authors point out in their review, Jeffreys immediately realized, however, that his multivariate rule does not generally produce sensible answers and suggested ad hoc alternatives in virtually all the multiparameter examples he analyzed, leading to a plethora of "Jeffreys priors" in the sense that they were proposed by him, although they do not follow from his general rule. Moreover, as all Bayesians in his time, Jeffreys was working under the assumption that a *unique* objective prior would be appropriate for all inference problems within a multiparameter model. Stein (1959) paradox already suggested that this could not possibly be true, but it was the discovery of the marginalization paradoxes (Dawid, Stone and Zidek, 1973) what definitely established this as a fact, while the reference priors (Bernardo, 1979) provided the first solution to the problem thus created.

José M. Bernardo is Professor of Statistics at the Facultad de Matemáticas, Universidad de Valencia, 46100-Burjassot, Valencia, Spain (e-mail: jose.m.bernardo@uv.es).