The Bayesian formalism deals elegantly with (b) but ignores (a). I am, of course, not suggesting that individual workers using Bayesian methods would necessarily ignore the issue, but the point is surely enough to refute any claim that the Bayesian formalism as normally presented is a comprehensive one. This point is related to 6.

In summary, I think it important to keep these arguments in perspective. The key issue in many applied statistical problems is not the framework for formal inference but rather the judicious choice of questions asked and their formalization, the choice of model in particular.

Comment

Simon French

As usual Professor Lindley has provided us with much to ponder. In this article he has presented the Bayesian position with great clarity; and, as a Bayesian, I can find little quarrel with most of his words. Nonetheless, as a discussant I have a duty, and there are one or two comments that I might make ....

There are times when someone says something with which you disagree, but you are unsure of precisely why. One such time happened to me in a *viva voce* examination, and Professor Lindley was my examiner. He has been pressing me on my understanding of the Bayesian position, and—I think—I had been reciting my catechism well. Then we came to the subject of utility. I searched for words to capture my understanding, but those that came did not find favor until I mumbled something about probability and gambles between the best and worst consequences. Professor Lindley leant back in his chair. Utility was probability, he explained, in the sense that he now uses in Sections 4.1 and 4.2. This explanation discomforted me, but I was unsure why. In any case, it certainly did not seem the time to press the matter further. Perhaps now is.

Professor Lindley contrasts the Bayesian paradigm with the "Berkeley" largely on their differing uses of probability. To me the difference seems more fundamental. In all analyses we build models. Some models are purely descriptive: they describe the world outside the person who "commissions" the analysis—call him or her the scientist. Other models are normative: included in them are a representation of the scientist's judgments. Moreover, in normative modeling the intention in representing the scientist's judgments is not to describe them as they are but to represent them in an idealized way (French, 1986; Bell, Raiffa and Tversky, 1988; and Phillips, 1984). For instance, in reality few people's beliefs are fully coherent, but we represent them in a Bayesian analysis by idealized, coherent probabilities. The purpose of doing so is to provide the scientist with guidance on how his or her judgments *should* evolve. If you like, the idealized behavior of the scientist built into the analysis is offered as a "role-model," suggesting how the scientist's views should change in the light of his or her data.

For me the difference between the Bayesian and Berkeley schools of inference lies in the unwillingness of the latter to use normative modeling. Berkeley statistics strives to be objective by leaving the judgments of the scientist out of the model. That Berkeley statistics ultimately fails in its attempt to be objective is a point on which I know Professor Lindley and I are in total agreement, and also one that I shall not pursue here. Rather let me return to the question of utility and its standing *vis à vis* probability.

Scientists are not the only people to have need of normative modeling. Decision makers, as Professor Lindley notes, also can benefit from its support. Normative modeling techniques can be used to guide and advise anyone who needs to express and act upon judgment in a formalized manner. Focusing on decision makers, the great strength of the Bayesian approach to decision analysis is that it separates the modeling of belief and uncertainty from the modeling of preference and value judgments, only drawing these two strands together in a coherent fashion at the end of the analysis. Now preference can exist independently of beliefs and uncertainty: I do not need to consider gambles to know that I prefer bananas to oranges. Since preference has an independent existence from beliefs and uncertainty, its modeling need not lead to a derived measurement system founded upon probability theory. There are ways of modeling preference in the complete absence of uncertainty: ordinal value functions, multi-attribute value functions, value difference functions, etc. None of these

Simon French is Professor of Operational Research and Information Systems, School of Computer Studies, University of Leeds, Leeds LS2 9JT, England.

need be assessed using probability and all can help in (the few) decision analyses where uncertainty is (effectively) absent. Only with the introduction of uncertainty does one need to assess attitude to risk and so introduce utility functions.

In presenting the theory underlying decision analysis, it is possible to keep the modeling of preference separate from the modeling of belief. Only at the end of the process, when one needs to combine such judgments to guide the choice of action do they come together (French, 1986). It is no quirk that in (Section 4.1, equation (1))

$$u(d, x) = \sum_y u(d, y)p(y \mid d, x)$$

$u(\cdot, \cdot)$ is an interval measurement unique to a positive affine transformation, whereas $p(\cdot \mid \cdot, \cdot)$ is an absolute measurement unique on the interval [0, 1]. Their different uniqueness properties derive from the fact they model different types of judgment.

I make this point not just to answer a question posed long ago by an examiner, but because utility has long been the poor relation of probability in the eyes of Bayesian statisticians. In one sense this is reasonable: statistics is essentially about the rational modification of belief in the light of data. However, as argued so cogently by Professor Lindley, as soon as one has need of a loss function then utility must enter the analysis. Because in statistical texts utility is so often treated as an add-on to subjective probability, the assumptions underlying its entrance into an analysis are often poorly appreciated. For instance, Professor Lindley writes (Section 4.2):

> "It is usually supposed by the adherents of utility theory that $L(d, y) \doteqdot \max_d u(d, y) - u(d, y)$, the difference between the best decision for y and that for the decision selected."

The presence of the minus sign in the supposed definition of $L(\cdot, \cdot)$ is important. It implies that the utility function not only represents attitude to risk but also strength of preference. The simultaneous representation of attitude to risk and strength of preference requires assumptions over and above those sketched by Professor Lindley in Section 4.1 (French, 1986, Chapter 9). These extra assumptions stand in relation to utility models in the same way that assumptions of probabilistic independence stand in relation to probability models. No statistician would be so cavalier as to forget to discuss the independence assumptions of a probability model; surely similar assumptions of any utility model deserve equal attention?

I suggested above that normative modeling guided the user by suggesting the way in which his or her judgments should evolve. If this is accepted, then Professor Lindley's analogy of the role played by probability theory in the assessment of belief and that played by Euclidean geometry in the mapping of the earth (Section 6.1) can, I believe, be taken too far. The earth is not a perfect sphere. Cartographers recognize this and seek to capture its "imperfections" faithfully in their maps.

An individual's beliefs are seldom perfectly coherent, yet a decision analyst would not seek to capture these imperfections in the probability model. Rather incoherencies would be referred back for reflection and revision. Much of the purpose of assessing subjective probabilities is to help an individual explore his or her beliefs, recognize any incoherence and revise some of them to make all his or her beliefs cohere. The analogy that Professor Lindley draws misses this. Cartographers do not seek to smooth off the hills and valleys; decision analysts do.

Taking this point further, Section 6.2 introduces "Your true probability $\pi(A)$" and contrasts this with its "direct measurement $p(A)$." My problem here is that I do not believe that individuals have "true" probabilities lying within them, which with suitable introspection and measurement can be dug out. Rather I believe in a more evolutionary view in which many of the individual's judgments are constructed through the interactive procedures of probability assessment. I do not have a probability for rain on the 14th July 1999 in Cleethorpes lying inside my mind waiting for discovery through introspection. If asked for my belief in such an event, I can bring together various pieces of knowledge and, with the support of the coherence of probability theory, can construct my subjective probability for a wet day in Cleethorpes.

Given this, I believe that the interpretation of the models investigated in Sections 6.1 to 6.8 needs much care. Nonetheless, I do believe that the paper makes a very real and important contribution here, because these models offer much to guide a decision analyst in devising elicitation strategies. (Ravinder, Kleinmuntz and Dyer, 1988, have carried out some related work.)

Statisticians, whatever school of inference they espouse, use parameters in their models and rightly so: the introduction of parameters make many modeling problems tractable. But what is the conceptual status of a parameter? For many Bayesians, de Finetti's seminal work on exchangeability together with much recent activity suggest an answer. Parameters are mathematical constructs within a model that encode exchangeability judgments. Professor Lindley's position on this is not entirely clear. In Section 2.2, he points to an interpretation based upon exchangeability; yet at other places he refers to "the practical reality ... of the parameter space" (Section 1.3), the "estimate $p(\theta \mid x)$" (Section 3.5), "the true value" (Section 3.6), the lack of a "qualitative difference between $p(x \mid \theta)$ and $\pi(\theta)$." In all these latter phrases there is the suggestion that θ has the same sort of reality as x, the observation.

This article has served to put into sharp contrast the Bayesian and Berkeley schools of statistics. Perhaps it is appropriate to close by remarking on a point of agreement between them.

By and large, all statisticians agree on the use of probability to model uncertainty. Perhaps we should unite on this agreement and look outside mainstream statistics. There we would notice a growth industry in *ad hoc* uncertainty modeling: fuzzy sets, possibility theory, varieties of belief representations, inexact logics, ... While we debate the niceties of priors versus sample spaces, there are many out there developing alternatives to our tools for inference and decision. Moreover, their alternatives, despite so many flaws obvious to us, are apparently far more attractive to those who award research and development funds. Many projects are building decision support systems and inference engines with what I can only describe as "inbuilt irrationality." Is it right that we stand idly by, waiting for their comeuppance? Professor Lindley is one of the few explaining carefully and patiently the flaws of these alternatives to probability modeling. It might be wise for us to forget, at least for the time being, some of the disagreements within statistics and put our energies into the wider debate of the value or otherwise of nonprobabilistic modeling of uncertainty.

ADDITIONAL REFERENCES

BELL, D. E., RAIFFA, H. and TVERSKY, A., eds. (1988). *Decision-Making: Descriptive, Normative and Prescriptive Interactions.* Cambridge Univ. Press, New York.

FRENCH, S. (1986). *Decision Theory: An Introduction to the Mathematics of Rationality.* Ellis Horwood, Chichester.

PHILLIPS, L. D. (1984). A theory of requisite decision models. *Acta Psychologica* **56** 29–48.

RAVINDER, H. V., KLEINMUNTZ, D. N. and DYER, J. S. (1988). The reliability of subjective probabilities obtained through decomposition. *Management Sci.* **34** 186–199.

Comment

Joseph B. Kadane

I want to supplement Lindley's admirable overview of Bayesian Statistics with some references and speculations about how modern computing may both influence Bayesian thought and be useful in accomplishing the agenda that Lindley, and before him Savage and others, have set out. The simplest Bayesian analyses, using exponential family likelihoods and stated priors in the conjugate form, do not require computing at all. Raiffa and Schlaifer (1961) give a still rather complete treatment of the computation of posterior distributions under these conditions. Modern Bayesian thought goes beyond these ideas in several respects. The important dimensions of generalization are:

a) The prior may not be stated, but may instead have to be elicited.
b) The likelihood may not be in the exponential family, or the prior may not conjugate with it.
c) The problem may not be the computation of a posterior distribution (or some functional of it) but rather a design problem.
d) Robustness may be of special concern.

I give some brief comments on each in turn.

Joseph B. Kadane is Leonard J. Savage Professor of Statistics and Social Sciences. His mailing address is Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.

1. ELICITATION

The idea of elicitation is to discover a prior that models the user's opinions well. Unfortunately this very important problem has not received the attention it deserves from the Bayesian computing community. For example, in Goel's (1988) survey of Bayesian programs, only two of the more than thirty listed concern elicitation, and neither of those was ready to be released. Nonetheless, this is a natural area for computation, particularly of the interactive sort. An early attempt of my own is given in Kadane, Dickey, Winkler, Smith and Peters (1980). For some more recent work in elicitation see Chaloner and Duncan (1983) and Gavaskar (1988). A very interesting recent work by DuMouchel (1988) uses graphical methods in the elicitation of a generalized ANOVA model.

As I have already remarked, I consider elicitation to be a very fruitful area for future work. One would think that the flexibility offered by modern devices such as mice would be useful in permitting users to express their views. While to date all the work reviewed here has assumed a given, known likelihood function, future elicitation work will, I believe, deal with the fact that likelihoods, as well as priors, are subjective and hence subject to elicitation (Bayarri, DeGroot and Kadane, 1988). Perhaps Lindley's work reported here will be the basis for future computer work in elicitation.