Discussion

Arthur Schleifer Jr., Dartmouth College

I should first like to congratulate Professors Roberts and Tiao for their excellent papers. Bayesian statistics owes a great debt to expositors like Harry Roberts whose paper is a clear, concise and well-balanced summary of many difficult and controversial problems; and George Tiao's paper, along with some of his previous publications, should appeal even to those who have not been attracted to Bayesian statistics by idealogical arguments, for it dramatically shows the power of the Bayesian approach in handling departures from the idealized assumptions of classical statistics. Although I am in basic agreement with almost every point made by Professors Roberts and Tiao, I shall devote this discussion to a difference of opinion regarding emphasis which in no way dampens my high regard for both papers.

In a session on Bayesian inference, I find it a bit surprising that neither paper attempts to distinguish formally between inference and decision theory. I must confess a certain uneasiness in trying to ascertain where the dividing line between these two areas lies, but I think the task is worthwhile because it forces us to consider what sorts of assessments of the object of uncertainty are useful.

In particular I should like to question the role in Bayesian inference of certain summary measures of posterior probability distributions to which Professor Roberts makes reference--for example, Bayesian point estimates, credible intervals, and the Bayesian analogue of significance tests. I shall argue that these measures either sacrifice experimental information or fail to process that information sufficiently, and that a tendency to summarize experimental results in terms of these measures can result in reports, such as one given in Professor Tiao's paper, which fail to serve the needs of the decision-maker as well as they might. Professor Tiao's example was, I am sure, intended to illustrate the application of a technique, and not to indicate how inferences should be made in the context of a real problem; but the possibility of misinterpretation by the casual reader is very real, and it is against this possible misinterpretation that I wish to argue.

Problems to which Bayesian analysis is applicable can be partitioned into an <u>experimental</u> phase and a <u>decision-making</u> phase, the output of the experimenter serving as input for the decision maker. No matter how "purely scientific" an experimenter may be, he must recognize that the report of his experimental results may some day be used in a decision context. At the other end of the spectrum, a rational decision maker, when acting under uncertainty, must take account of whatever objective experimental evidence is available and whatever subjective or informal evidence he may have acquired elsewhere.

Given these two phases of problem analysis, we may specify that the output of Bayesian decision theory is a <u>course of action</u>, while the output of Bayesian inference is a <u>report</u> which serves as input for the decision-maker. It is generally agreed that such a report should <u>summarize</u> the experimental data and <u>process</u> the data in the form most useful to the decision-maker, but that <u>under no circumstances should the report sacrifice information</u>.

The prescription for fulfilling these goals depends on the gap between experimenter and decision-maker. If the experimenter cannot even anticipate the decision-maker's data-generating model, then it is best simply to report the raw experimental data. If the experimenter <u>can</u> specify a data-generating model, but is unable to anticipate the decision-maker's prior distribution and the economic consequences of each possible course of action, then the raw data, sufficient statistics, likelihood function, or distribution posterior to a diffuse prior will each convey all of the experimental information. A strong case can be made for choosing the posterior distribution as the best summary of the experimental information: it is easily combined via Bayes' theorem with the decision-maker's prior distribution provided the latter can be expressed as a pseudo-sample, and it may convey sufficient information by itself to permit the decision-maker to forego the sometimes painful process of assessing his prior.

In this case any summary measures of the reported posterior distribution run the risk of losing information by preventing the decision-maker from incorporating his own prior information. By way of illustration, consider Professor Tiao's example for the comparison of two variances. The data happen to consist of measurements made by an inexperienced analyst A_1 and an experienced analyst A_2 in performing a chemical assay. As the man responsible for making decisions about A_1's competence, I might be very interested in <u>a</u> posterior probability that $V > 1$ -- but not in <u>the</u> posterior probability that $V > 1$ given in Professor Tiao's paper. I would certainly want <u>my</u> probability to reflect what I already know of the relative abilities of A_1 and A_2, of the amount of experience A_1 has already had, of my knowledge of learning rates, etc., but the posterior probability that $V > 1$ given by Professor Tiao, whether conditional on β or marginal, does not permit me to incorporate this information. What is really needed for my purpose is <u>the marginal distribution of V</u> posterior to a diffuse prior.

Having considered the case in which the experimenter knows only the data-generating model which the decision-maker will use, let us turn to the case in which the experimenter, before conducting the experiment, possesses all of the prior and economic information of the decision-maker. In this case he may indeed calculate the distribution of the decision parameter posterior to both the experiment and the decision-maker's prior distribution, and, with knowledge of the economic structure of the problem, may even obtain summary measures of this distribution without sacrificing information. But I maintain that in this case a report of such measures does not constitute sufficient processing of the experimental data: a report of the <u>posterior expected utility</u> of each act is in keeping with our goal of processing the experimental results as far as is possible without sacrificing relevant information.

Summary measures of the posterior distribution <u>may</u> be relevant when the experimenter knows the prior distribution of the decision-maker and the functional form of the utility of each act with respect to the decision variable, but does not know the parameters of these functions. Such cases, however, are too rare to justify these measures.

Surely these summary measures are important because of their relation to the analogous classical point estimates, confidence intervals, and significance tests; they are also interesting conceptually and useful in the teaching of Bayesian inference and decision theory. But I think that their importance among the tools of the trade has been overemphasized.