(3) It would be of interest to obtain an expansion of the form

$$i^X(\theta) - i^T(\theta) = \alpha + \frac{\beta}{n} + \cdots.$$

We know the expression for α and its geometric interpretation. What about β?

(4) I believe that the choice of a prior distribution is governed by the nature of the parameter and previous knowledge (though vague) about it and should not depend on what experiment is conducted to have further information on it. Jeffreys' invariant prior may have nice properties but it seems to depend on how observations are generated, which may not be acceptable to Bayesians.

ADDITIONAL REFERENCES

RAO, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Cambridge Philos. Soc.* **44** 50–57.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

Comment

N. Reid and D. A. S. Fraser

We congratulate Professor Kass on a very clear and interesting account of the role of differential geometry in asymptotic inference. In particular, his discussion of information loss and recovery through conditioning, and the geometric interpretation of this, adds substantially to the long-standing discussion initiated in Fisher's early work.

The use and implications of conditional analysis are central to the topics in the paper. In this discussion, we expand a little on arguments for and justifications of conditioning, and the use of geometric methods to motivate this.

In the setting discussed in Section 3.1, we can write

$$(1) \qquad p_Y(y \mid \theta) = p_{T|A}(t \mid a, \theta)p_A(a)$$

where $Y = (T, A)$ is sufficient, A is ancillary, and the Jacobian has been absorbed into the support differentials. This factorization suggests, as the paper indicates, that inference about θ may be based on the conditional distribution of T given A, without loss of information about θ. Section 3.1.1 gives formal clarity to Fisher's general analysis of information loss and is valuable in giving a precise interpretation of the phrase "without loss of information about θ."

Other arguments can also provide some interpretation of the phrase above. For example the likelihood function obtained from the conditional distribution is the same as the likelihood function from the distribution of the full data Y. Another motivation for conditioning on A when the factorization in (1) holds is that the variability in the outcome that is described by the marginal distribution of A is irrelevant for inference about θ; this is an underlying theme in Fisher's early work expanded in Fisher (1961) and is very clearly presented in the weighing machine example of Cox (1958). Fisher frequently used the term "relevant subset" to refer to the set of sample points having the observed value for the ancillary statistic. However, it seems clear that he attached additional meaning to the term, derived from the physical context from which the statistical problem arose. Indeed, this additional interpretation may well have been primary in Fisher's interpretation of conditioning and the definition of the correct probabilities to use in applications. There does seem to be no fully satisfactory formalization of such "relevant subsets" based on the statistical model alone. The derivation of the Likelihood Principle from the Conditionality Principle discussed in Evans, Fraser and Monette (1986) bears on this.

Most discussions of conditioning are motivated by a few very compelling examples. Subsequent attempts to formalize the operating principle to enable extension to more realistic settings are widely divergent. One development, primarily initiated by Birnbaum (1962, 1972) and Basu (1959, 1964) (see also Buehler, 1982), isolates ancillarity as the essential feature; the discussion of this approach and its relation to Bayesian inference and the likelihood principle is well summarized in Berger and Wolpert (1985).

Another development of conditioning in Fraser (1968, 1979) extends and formalizes one aspect of

N. Reid is Professor in the Department of Statistics at the University of Toronto. Her mailing address is: Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S 1A1. D. A. S. Fraser is Professor in the Department of Mathematics at York University. His mailing address is: Department of Mathematics, York University, Downsview, Ontario, Canada M3J 1P3.