INFERENCE 185

Expression (16) suggests that it might be impossible to find a prior density that produces confidence limits having coverage error of order $O(n^{-3/2})$; see DiCiccio, Keller and Martin (1992).

Many of the likelihood adjustments and distributional corrections discussed in the paper can be viewed, at least to error of order $O_p(n^{-1})$, in terms of the quantities z_0 and a that arise in Efron's (1987) BC_a confidence limits. Efron defined $z_0 = \Phi^{-1}\{\operatorname{pr}(\hat{\psi} \leq \psi_0)\}$, and a is related to the skewness of the score function; both z_0 and a are of order $O(n^{-1/2})$. In the setting of Section 4.2, DiCiccio and Efron (1992) and Efron (1993) showed that $E(r_p) = -z_0 + O(n^{-1})$ and that $r_p + z_0$ has the standard normal distribution to error of order $O(n^{-1})$. Moreover,

$$E\{l_p'(\psi)\} = (a-z_0)\{-l_p''(\psi)\}^{1/2} + O(n^{-1})$$

and

$$l_c(\psi) = l_p(\psi) - (a - z_0) \{-l_p''(\psi)\}^{1/2} + O(n^{-1}).$$

As many authors have noted, adjustment of the log profile likelihood function $l_p(\psi)$ reduces the bias of the profile score. Also, $E(r_c) = -a + O(n^{-1/2})$, and $r_c + a$ has the standard normal distribution to error of order $O(n^{-1})$. Further details are given in DiCiccio and Efron (1995).

ACKNOWLEDGMENTS

The first author's research is supported by NSF Grant DMS-93-05547. This is paper BU-1283-M in the Biometrics Unit, Cornell University, Ithaca, New York 14853. The third author's research is supported by NIH Grant No. RO1-CA61120.

Comment

A. P. Dawid and C. Goutis

Nancy Reid has presented a clear and valuable overview of the uses of conditioning, and of associated techniques of analysis. We wish to focus on some difficulties which can arise from too uncritical an attitude to conditional inference.

It is implicit in Reid's account, as in most others, that the goal of conditional inference has been achieved when we have identified the appropriate conditional "frame of reference" (Dawid, 1991). From that point on, it is implied, we should be free to use any favourite method of inference within that new frame. However, a more thorough-going analysis casts doubt on this assumption. This doubt may be evidenced in several related ways.

First there is the problem of nonuniqueness of (maximal) ancillary statistics, and the consequent arbitrariness, in general, of the conditional frame of reference. The collected works of Basu (1988), which deal thoroughly with these topics, should be required reading for anyone contemplating conditional inference. For example, if (X_i, Y_i) have a bivariate normal distribution with known vari-

Philip Dawid is Professor of Statistics and Costas Goutis is Lecturer in Statistics, Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, United Kingdom. ances and unknown correlation ρ , each of $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ is ancillary, and inference conditional on either appears equally justified. We cannot, however, condition on both, since (\mathbf{X}, \mathbf{Y}) reproduces the whole sample.

Next, there is Birnbaum's (1962) celebrated demonstration that acceptance of both the sufficiency and conditionality principles demands acceptance of the likelihood principle—and is thus incompatible with any method of inference which does not respect that principle. A much weaker version of this argument and conclusion, which nevertheless implies the irrelevance of optional stopping and is hence incompatible with many common forms of inference, is given by Dawid (1986).

Then there is the "conflict between conditioning and power" mentioned in Section 6.2. A concrete example, based on Cox (1958a), is analysed in Dawid (1983, pages 99–100). In a problem with point null and alternative hypotheses, and a simple experimental ancillary, the rule "use the likelihood ratio test with size $\alpha=0.05$," if applied conditionally on the ancillary, does not agree with any unconditional likelihood ratio test and is thus less powerful than the overall 0.05-level test (which has differing conditional α -levels). However, the Neyman–Pearson lemma, which simply requires use of some likelihood ratio test, can nonetheless be ap-