The Annals of Statistics
1983, Vol. 11, No. 2, 368–369

DISCUSSION

JAMES O. BERGER

Purdue University

The authors are to be congratulated for producing a significant contribution to mathematical statistics. Before discussing their specific contribution, a few general background comments are in order.

The impetus for constructing improved simultaneous estimators has generally come from two directions: (i) the decision-theoretic approach involving production of estimators dominating (in terms of risk) "usual" estimators which are inadmissible in higher dimensions, and (ii) the Bayesian or empirical Bayesian approach of taking advantage of prior information about the unknown parameters (often information concerning a plausible structure or "model" for these parameters) to produce substantially better estimators tailored to the supposed prior information. It is generally the case that the Bayesian or empirical Bayesian approach is of more practical interest, especially when the unknown parameters θ_i are thought to have some common structure, in that very substantial improvement over standard estimators is usually possible only in such situations. As with any approach involving a more restrictive model, however, one runs the risk of worse performance if the presumed structure is actually not present.

The decision theoretic approach does not typically provide estimators offering as substantial improvement, partly because one is constrained to seek dominating estimators, i.e., estimators never worse than standard estimators, in terms of risk. This dominance is attractive to many statisticians, however, since there then need be no fear of disastrous inferences if, say, a presumed structure for the θ_i turns out not to be present. (One sour note here is that risk dominance of an estimator may be very dependent on the assumed loss function, a quantity hard to know or specify in practice.)

The interaction between the two approaches has been very beneficial for the subject of improved simultaneous estimation of normal means. Bayesian and empirical Bayesian ideas have indicated what basic types of estimators are needed if practical gains are to be significant; decision theoretic ideas have shown how to fine tune or modify these estimators so as to achieve robustness with respect to errors in the specification of the prior or structure for the θ_i. It should be emphasized that progress, especially on the mathematically difficult decision-theoretic side, usually was made in small steps, with many of the intermediate steps being of negligible practical utility but contributing to more useful later results.

The discrete situation discussed in this paper has been much less extensively studied than the normal mean situation, especially from the decision-theoretic side. Indeed, only very recently have estimators dominating the usual estimator been found which allow utilization of prior information through the choice of various shrinkage patterns. The theoretical techniques developed by the authors for establishing dominance results for such estimators are truly impressive, and provide a significant advancement in theory. It seems clear, however, that much remains to be done. In particular, the estimators in the paper do not appear to match the performance of the Bayes or empirical Bayes (but non-dominating) estimators considered in such articles as Albert (1981) and Hudson and Tsui (1981), or the normal theory estimators that can be employed by making an approximate transformation to normality (such as δ^M as evaluated in Table 2). It is possible that risk domination is too severe a restriction in this problem and should be abandoned. Even if risk domination is relaxed, however, the theoretical techniques of the article may well be useful, as in Berger (1982a) for the normal situation.

Received October 1982.

One surprise provided by the article is the comparatively good performance (among dominating estimators) of δ^2, the estimator shifting towards the minimum of the Poisson observations. Such an estimator is not a typical empirical Bayes estimator, and its comparatively good performance indicates that such estimators may deserve more study. (One of the useful features of the decision-theoretic approach is that it often does produce "surprises"—witness the original paper of Stein (1956), which to a large extent got the whole subject rolling.)

Perhaps the most disturbing feature of the article, and indeed the decision-theoretic approach to this problem in general, is the considerable dependence of the results on the loss function. In the normal means problem, the use of a location loss is very reasonable, and the common assumption of squared error loss (or weighted squared error loss) is convenient and seems to be moderately unimportant as far as dominance goes (though the weights can have a substantial effect). In the Poisson problem, on the other hand, the choice of an appropriate loss is much less clear, and even the losses L_m considered in the paper seem to give quite different results for different m. What we need are estimators which work well (even if they are not completely dominating) for a variety of plausible loss functions.

It should be mentioned that, if properly constructed, improved simultaneous estimators can be reasonable from a conditional (Bayesian, say) perspective (c.f. Berger, 1982b). Hence the work is not completely uninteresting to a conditionalist, even though frequentist measures such as risk are used.

In conclusion, the present paper provides us with substantial new theoretical techniques for proving dominance and develops some interesting estimators which suggest useful possibilities to explore (and which may be useful in special situations). Substantial practical advances, however, are more likely to follow from development of Bayes or empirical Bayes estimators tailored to specific types of prior information or parameter structures. The theoretical techniques in this paper should prove useful even in such developments as tools for establishing Bayesian robustness or insensitivity to prior assumptions.

REFERENCES

ALBERT, J. (1981). Simultaneous estimation of Poisson means. *J. Multivariate Anal.* **11** 400–417.

BERGER, J. (1982a). Estimation in continuous exponential families: Bayesian estimation subject to risk restrictions and inadmissibility results. In *Statistical Decision Theory and Related Topics III*, S. S. Gupta and J. Berger (eds.). Academic, New York.

BERGER, J. (1982b). Bayesian robustness and the Stein effect. *J. Amer. Statist. Assoc.* **77** 358–368.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907