Of course, these arguments prove nothing about admissibility but do suggest that the necessity for the known mean of the V_i's is not unreasonable.

School of Mathematics and Statistics
The University of Birmingham
P.O. Box 363
Birmingham B15 2TT
United Kingdom

B. Efron

Stanford University

Here is a slightly simpler version of Brown's nice paradox: the statistician observes $X \sim N_p(\mu, I)$, $p \geq 3$, and also an integer J that equals $j = 1, 2, 3, \dots, p$ with probability $1/p$, independently of X. It is desired to estimate μ_j with squared-error loss. Then J is ancillary, and conditional on $J = j$ the obvious estimate $d_0(X, j) = X_j$ is admissible and minimax. Unconditionally, however, the Jth coordinate of the James–Stein estimate,

$$d_1(X, J) = \left[1 - (p-2)/\|X\|^2\right] X_J,$$

dominates $d_0(X, J)$, with $E[d_1(X, J) - \mu_J]^2 < E[d_0(X, J) - \mu_J]^2$ for all vectors μ.

In other words, Brown has restated Stein's paradox, that d_1 dominates d_0 in terms of total squared error loss, in an interesting way that casts some doubt on the ancillarity principle.

[The example above does not look much like Brown's regression paradox, but we can fix things up by supposing that given $J = j$ the statistician also observes $X_0 \sim N(\alpha + \mu_j, 1)$, independent of $X \sim N_p(\mu, I)$, the goal now being to estimate α with squared-error loss. Then $\hat{\alpha}_1 = X_0 - d_1(X, J)$ dominates $\hat{\alpha}_0 = X_0 - d_0(X, J)$ unconditionally but not conditionally. This situation might arise if X_j was the placebo response of patient j on some physiological scale and X_0 was patient j's response when given a treatment of interest; we placebo-test p patients and then choose one at random to receive the treatment.]

Why do we intuitively accept the ancillarity principle in Cox's example, Section 5, but doubt it in the example above, or in Brown's regression paradoxes? I believe that the answer has more to do with single versus multiple inference than with hypothesis testing versus estimation.

Notice that $d_0(X, j)$ disregards all of the data except X_j. There is nothing in the ancillarity principle to justify this. All that ancillarity says is that we should do our probability calculations conditional on $J = j$. In Cox's example on the other hand, the conditional solution makes use of all the data and the ancillarity principle works fine.

Even when the choice $J = j$ is totally nonrandom it is not obvious that d_0 is preferable to d_1. The real question is whether or not the ensemble estimation gains offered by d_1 are relevant to the specific problem of estimating μ_j.

Carl Morris and I worried about this a lot in our 1971 and 1972 papers, and also in the specific examples of 1975. Our hard-working 18 baseball players were offered as a simplified test case for thinking about the trade-offs between d_0 and d_1; see also Section 8 of Efron (1982).

REFERENCES

EFRON, B. (1982). Maximum likelihood and decision theory. *Ann. Statist.* **10** 340–356.

EFRON, B. and MORRIS, C. (1971). Limiting the risk of Bayes and empirical Bayes estimators, I: The Bayes case. *J. Amer. Statist. Assoc.* **66** 801–815.

EFRON, B. and MORRIS, C. (1972). Limiting the risk of Bayes and Empirical Bayes estimators, II: The empirical Bayes case. *J. Amer. Statist. Assoc.* **67** 130–139.

EFRON, B. and MORRIS, C. (1975). Data analysis using Stein's estimator and its generalizations. *J. Amer. Statist. Assoc.* **70** 311–319.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

D. A. S. FRASER AND N. REID

York University and University of Toronto

1. Introduction. Professor Brown has presented a comprehensive discussion of multiple regression in relation to admissibility and the ancillarity principle. He concludes that there is a paradox: That the results with multiple regression contradict "the widely held notion that statistical inference in the presence of ancillary statistics should be independent of the distribution of these ancillary statistics." The reader thus receives the impression that there is something wrong or inappropriate with conditional inference. The basic assumption of conditional inference is that only the conditional model is examined and that information from the marginal model is ignored. This is not a "notion" that inference "should" be independent of the marginal model as interpreted by Professor Brown, but that inference should not use or make reference to that model.

The technical point then is that there is a conflict between conditional methods and classical optimality criteria. We feel that this should be no surprise, let alone paradox. In Section 5 we present a simple example that also illustrates the conflict.

Our broader viewpoint is that the familiar optimality criteria of statistics are in fact in conflict with scientific principles and that this provides the explanation for the issues raised in the paper; see Section 2.

In a concluding Section 6, we argue that conditional methods are close to the core of the scientific method, and note that conditional inference from both a theoretical and pragmatic orientation is a recently active area of research and presents exciting possibilities for research development.

Standard statistical theory uses a range of optimality criteria, such as maximum power for a test at a given size α, minimum length for a confidence