*The Annals of Statistics*
1975, Vol. 3, No. 2, 448-450

# ADMISSIBILITY OF THE BEST INVARIANT ESTIMATOR OF ONE CO-ORDINATE OF A LOCATION VECTOR[1]

BY STEPHEN L. PORTNOY

*Harvard University*

In 1960, Charles Stein conjectured that the best invariant estimate of a single co-ordinate of a $p$-dimensional location parameter would be admissible if $p \leqq 3$ but inadmissible if $p \geqq 4$. This paper presents a class of examples which supports Stein's conjecture.

In [4], Charles Stein made some conjectures about the estimation of one or two co-ordinates of a multivariate location parameter with quadratic loss. In particular, he conjectured that the best invariant estimator of a single co-ordinate of a $p$-dimensional location parameter would generally be admissible if $p \leqq 3$, but would not generally be admissible if $p \geqq 4$. In that paper, the admissibility results were proven for $p = 1$ and $p = 2$; and Stein noted that a pathological example due to Blackwell [3] could be restated as an example for $p = 4$ where the best invariant estimate is inadmissible.

L. Brown and J. Berger (see [1] and [2]) have recently considered this problem in considerable generality and essentially answered Stein's conjecture in the affirmative. They reduce admissibility problems, at least approximately, to the existence of solutions to certain partial differential equations; and their inadmissibility results make use of known facts about these equations. However, the admissibility results in [2] use the classical method of exhibiting an appropriate sequence of prior densities for which the Bayes risk converges sufficiently quickly to the expected risk. In this case, however, the calculations are extremely lengthy and complicated. The present paper reduces an admissibility problem to a difference equation and then uses the probabilistic techniques given in [5] to solve the admissibility problem. To my knowledge, these techniques are novel and may eventually prove very useful in considering more realistic admissibility problems.

In particular, consider the following class of location parameter problems in $p$ dimensions: Let $\theta_1 \in R$, $\theta_2 \in R^{p-1}$, and let $\theta = (\theta_1, \theta_2)$ be the location parameter. Let $Q$ denote a probability measure on $R^{p-1}$ and let $X \sim Q$. Let $Y$ be a random variable, $Z$ a random vector in $R^{p-1}$, and define the distribution of the observation $(Y, Z)$ as follows under $\theta$: with probability $\frac{1}{2}$, $Y = \theta_1 - 1$ and $Z = \theta_2 + X$; and with probability $\frac{1}{2}$, $Y = \theta_1 + 1$ and $Z = \theta_2$. We want to estimate $\theta_1$ with

Received July 1973; revised March 1974.

[1] This work was facilitated by grants GS-32327X and GP-31003X from the National Science Foundation.

*AMS* 1970 *subject classifications*. 62C15, 62F10.

*Key words and phrases*. Admissibility, location invariance, estimation with nuisance parameters.

squared error loss. It is clear that, independent of $Q$, the best invariant estimator exists and is $\phi_0(Y, Z) = Y$.

Now let $S = \{(\theta_1, \theta_2) : \theta_1$ is an even integer$\}$. Then $Y$ is an odd integer with probability one, and it can be shown that $\phi_0$ is admissible if and only if it is admissible on $S$. Thus, it suffices to consider estimators of the form

$$\phi(y, z) = y - f_n(z) \qquad \text{where} \quad y = 2n - 1$$

where $\{f_n : n = 0, \pm 1, \pm 2, \cdots\}$ is an arbitrary sequence of measurable functions from $R^{p-1}$ to $R$. For $\theta_1 = 2n$, the risk of such $\phi$ is

$$\begin{aligned} R(\phi, \theta) &= E_\theta(Y - f_N(Z) - \theta_1)^2 \quad (\text{where } \; Y = 2N - 1) \\ &= E_\theta\{(Y - \theta_1)^2 - 2(Y - \theta_1)f_N(Z) + f_N{}^2(Z)\} \\ &= 1 - (2n + 1 - 2n)f_{n+1}(\theta_2) + \tfrac{1}{2}f_{n+1}^2(\theta_2) \\ &\quad - (2n - 1 - 2n)E_Q f_n(\theta_2 + X) + \tfrac{1}{2}E_Q f_n{}^2(\theta_2 + X) \\ &= 1 + E_Q[f_n(\theta_2 + X) + \tfrac{1}{2}f_n{}^2(\theta_2 + X)] - f_{n+1}(\theta_2) + \tfrac{1}{2}f_{n+1}^2(\theta_2)\,. \end{aligned} \tag{1}$$

Now let $T$ be the linear operator defined for $f\colon R^{p-1} \to R$ by

$$(Tf)(x) = E_Q f(x + X) = E_{Q_x} f(X) \tag{2}$$

where $Q_x$ is the distribution of $x + X$. Thus, $T$ is an expectation operator for each $x$. Therefore, from (1) and (2) we see that $\phi_0$ is inadmissible if and only if there is a sequence $\{f_n : n = 0, \pm 1, \pm 2, \cdots\}$ with some $f_n$ not zero under every $Q_\theta$ such that

$$f_{n+1}(x) - \tfrac{1}{2}f_{n+1}^2(x) \geqq T(f_n + \tfrac{1}{2}f_n{}^2)(x)\,. \tag{3}$$

Note that strict inequality is not required in (3) because the loss function is strictly convex. Thus, if two estimators have the same risk function but are not equal almost surely under every $Q_\theta$, a convex combination of them will have strictly smaller risk (under some $Q_\theta$).

Equation (3) is treated in detail in reference [5]. The basic results presented there are the following:

(a) If $Q$ is a lattice distribution, has finite second moment, and $p = 2$ or $p = 3$, then equation (3) has no nonzero solution; in particular, the best invariant estimator is admissible.

(b) If the random walk generated by the symmetrization of $Q$ is transient, then equation (3) has a nonzero solution. In particular, if $p \geqq 4$ (or if $Q$ is stable with index $\alpha < 1$ in $p = 2$ or stable with index $\alpha < 2$ in $p = 3$) then the best invariant estimator is inadmissible.

Some final remarks should be made:

(1) If $Q$ is not discrete, then $\phi_0$ is always inadmissible, since it can be changed at a single point to obtain improvement at a single parameter value. Thus, the discreteness of $Q$ is actually in the nature of a regularity condition imposed by the singularity of the distributions. However, the additional moment conditions

required for admissibility seem somewhat surprising, since $\phi_0$ itself has infinitely many moments.

(2) The results of [5] actually show that in general, if $Q$ has finite second moment and $p = 2$ or $p = 3$, then $\phi_0$ is almost admissible with respect to the measure consisting of counting measure on the first co-ordinate crossed with Lebesgue measure on the remaining $(p - 1)$ co-ordinates.

(3) It seems clear that the proofs in [5] could be extended to the case where $Q$ is not concentrated on a hyperplane, but is defined on a half-space bounded away from the point mass at $(\theta_1 + 1, \theta_2)$. However, the point mass does seem to be necessary in order to consider (3) in terms of a random walk. The general case would at best lead to an equation like (3) but with a different expectation operator also appearing on the left-hand side. Means of solving this equation are not immediately apparent.

## REFERENCES

[1] Berger, J. (1974a). Inadmissibility results for generalized Bayes estimators of coordinates of a location vector. Technical Report 352, Purdue Univ.

[2] Berger, J. (1974b). Admissibility results for generalized Bayes estimators of coordinates of a location vector. Technical Report 353, Purdue Univ.

[3] Blackwell, D. (1951). On the translation parameter problem for discrete variables. *Ann. Math. Statist.* **22** 393-399.

[4] James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press. 361-379.

[5] Portnoy, Stephen (1975). Transience and solvability of a non-linear diffusion equation. To appear in *Ann. Prob.* No. 3.

Department of Mathematics
University of Illinois
Urbana, Illinois 61801