The Annals of Statistics
1974, Vol. 2, No. 3, 553–557

ALTERNATIVE ESTIMATORS FOR THE SCALE PARAMETER OF THE EXPONENTIAL DISTRIBUTION WITH UNKNOWN LOCATION[1]

BY J. F. BREWSTER

University of Manitoba

Let $X_1, X_2, \cdots, X_n$ be independent observations from an exponential distribution with an unknown location-scale parameter (μ, σ). Let $\bar{X} = n^{-1} \sum X_i$ and $M = \min X_i$. Under squared error loss the best location-scale equivariant estimator of σ is $\bar{X} - M$, which agrees with the maximum likelihood estimator. Arnold (*J. Amer. Statist. Assoc.* **65** (1970) 1260–1264) and Zidek (*Ann. Statist.* **1** (1973) 264–278) have shown that $\bar{X} - M$ is inadmissible, but the dominating estimator which they produce is probably inadmissible as well. In this paper a "smoother" dominating procedure is presented, and the risk functions of the various alternatives are plotted. Similar results are obtained for strictly bowl-shaped loss functions.

1. Introduction. Let $X_1, X_2, \cdots, X_n$ be independent observations from an exponential distribution with unknown location-scale parameter (μ, σ), and consider the estimation of σ under squared error loss. That is, $L(\hat{\sigma}; \mu, \sigma) = \sigma^{-2}(\hat{\sigma} - \sigma)^2$. A sufficient statistic is $(M, \bar{X})$, where $M = \min X_i$ and $\bar{X} = n^{-1} \sum X_i$, and, for convenience, we let $S = \bar{X} - M$. The problem remains invariant under the location-scale group, $\mathscr{G}$, and any nonrandomized $\mathscr{G}$-equivariant estimator is of the form cS, for some $c > 0$. As $\mathscr{G}$ acts transitively on the parameter space, the risk function of such a procedure is constant, and the best $\mathscr{G}$-equivariant estimator is S, coinciding with the maximum likelihood estimator.

We consider the invariance of the problem, not under $\mathscr{G}$, but under the scale subgroup, $\mathscr{S}$. If we let $Z = MS^{-1}$, then any $\mathscr{S}$-equivariant estimator is of the form $\phi(Z)S$. In analogy with Stein's estimator (1964) for the variance of a normal distribution, Arnold (1970) and Zidek (1973) have produced an $\mathscr{S}$-equivariant estimator, $\phi^*(Z)S$, which has uniformly smaller risk than S. It is given by

$$\begin{aligned}(1.1)\qquad \phi^*(z) &= \min\{1, n(n+1)^{-1}(1+z)\} && z \geqq 0 \\ &= 1 && z < 0 .\end{aligned}$$

Arnold demonstrates the result by deriving the risk function of $\phi^*(Z)S$ and comparing it with that of S. Zidek's proof entails the examination of the conditional expected loss given the maximal invariant, Z. The latter proof illustrates

Received January 1973; revised July 1973.

[1] This research was supported in part by a grant from the National Research Council of Canada. Section 2 of this paper comprised a portion of the author's Ph. D. dissertation completed at the University of British Columbia.

AMS 1970 *subject classifications.* Primary 62C15; Secondary 62F10.

Key words and phrases. Exponential scale estimators, improving estimators, bowl-shaped loss functions.

one of two closely related techniques for improving estimators presented by Brewster and Zidek (1974), in a paper in which the analogous problem for the normal distribution is discussed. The estimator $\phi^{**}(Z)S$, which will be obtained here, illustrates the second technique.

In the case of the normal distribution the first method produces the estimator of Stein, and the second method produces a "smoother" estimator which is also generalized Bayes and admissible within the class of $\mathscr{S}$-equivariant estimators. One might hope that the same would be true of $\phi^{**}(Z)S$. Although ϕ^{**} is "smoother" than ϕ^*, it is unfortunately discontinuous at the origin, and is not generalized Bayes. In view of an example of Sacks (1963) regarding the exponential distribution, however, perhaps we are asking too much of a candidate for admissibility in this problem when we ask that it be generalized Bayes.

2. Construction of ϕ^{**}. As a first step in the construction of ϕ^{**} we demonstrate the inadmissibility of S by adapting a method of Brown (1968). Let

$$\Psi(x) = \frac{1-(1+x)^{-n}}{1-(1+x)^{-n-1}}, \qquad x > 0. \tag{2.1}$$

LEMMA 2.1. *For fixed $r > 0$, $\phi_r(Z)S$ has uniformly smaller risk than S, where*

$$\begin{aligned} \phi_r(z) &= \Psi(r) && 0 < z \leqq r \\ &= 1 && z \leqq 0 \text{ or } z > r. \end{aligned} \tag{2.2}$$

PROOF. Consider estimators of the form $\phi(Z)S$, where $\phi(z) \equiv c$, $0 < z \leqq r$, and $\phi(z) = 1$, otherwise. The risk function of such an estimator depends only on $\lambda = \mu\sigma^{-1}$, and is given by

$$E_{\lambda,1}([cS-1]^2 \mid Z \in (0, r]) \cdot P_{\lambda,1}(Z \in (0, r]) + E_{\lambda,1}([S-1]^2 \mid Z \notin (0, r]) \cdot P_{\lambda,1}(Z \notin (0, r]).$$

If c were allowed to depend on λ, then the optimum choice of c would minimize $E_{\lambda,1}([cS-1]^2 \mid Z \in (0, r])$. In other words,

$$c_\lambda = \frac{E_{\lambda,1}(S \mid Z \in (0, r])}{E_{\lambda,1}(S^2 \mid Z \in (0, r])}.$$

But

$$\sup_\lambda c_\lambda = c_0 = \Psi(r) < 1, \tag{2.3}$$

and since $E_\lambda([cS-1]^2 \mid Z \in (0, r])$ is a "strictly bowl-shaped" (in fact, convex) function of c, the proof is complete. □

Now select $0 < r' < r$. By repeating the previous argument, and noticing that $\Psi(r') < \Psi(r)$, we are able to conclude that $\phi_{r',r}(Z)S$ has uniformly smaller risk than $\phi_r(Z)S$, where

$$\begin{aligned} \phi_{r',r}(z) &= \Psi(r') && 0 < z \leqq r' \\ &= \Psi(r) && r' < z \leqq r \\ &= 1 && z \leqq 0 \text{ or } z > r. \end{aligned} \tag{2.4}$$

We can clearly continue to produce step-function estimators by selecting successively smaller constants. By noticing that the starting point, r, is arbitrary, we obtain the following theorem.

THEOREM 2.1. *The risk function of* $\phi^{**}(Z)S$ *is nowhere larger than that of* S, *where*

$$\phi^{**}(z) = \Psi(z) \qquad z > 0$$
$$= 1 \qquad z \leqq 0 . \tag{2.5}$$

PROOF. For each $i = 1, 2, \cdots$, select a finite partition of $[0, \infty)$ represented by $0 = r_{i0} < \cdots < r_{in_i} < \infty$, and a corresponding estimator $\phi^{(i)}(Z)S$, where

$$\phi^{(i)}(z) = \Psi(r_{ij}) \qquad r_{i,j-1} < z \leqq r_{ij}$$
$$= 1 \qquad \text{otherwise} .$$

Then, providing $\max_j |r_{ij} - r_{i,j-1}| \to 0$ and $r_{i,n_i} \to \infty$, the sequence $\phi^{(i)}$ will converge pointwise to ϕ^{**}.

Although we are, in general, unable to compare $\phi^{(i)}$ and $\phi^{(i')}$, $i \neq i'$, we do know that $\phi^{(i)}(Z)S$ has uniformly smaller risk than S, for all i. The proof is completed by applying Fatou's Lemma. □

REMARK. From Figure 1 it is apparent that $\phi^{**}(Z)S$ actually dominates S.

The following corollary is an obvious consequence of the method of proof in Theorem 2.1.

COROLLARY 2.1. *If* $\phi^{**}(z) \leqq \phi(z) \leqq 1$ *for all* z, *and* ϕ *is non-decreasing on* $(0, \infty)$, *then the risk function of* $\phi(Z)S$ *is nowhere larger than that of* S.

3. Extension to strictly bowl-shaped loss. Although we assumed squared error loss in Section 2, the specific form of the loss function played only a minor role in the proofs. Here, we consider the estimation of σ^m $(m > 0)$, and assume that $L(a; \mu, \sigma) = W(a\sigma^{-m})$ is a nonnegative $\mathscr{G}$-invariant loss function. We assume, in addition, that W is a strictly bowl-shaped function. In other words, W is strictly decreasing on $(0, u_0]$, and strictly increasing on $[u_0, \infty)$, for some $u_0 > 0$. As a consequence, W is differentiable almost everywhere, and we assume, whenever necessary for integrals involving W, that interchange of derivative and integral is permissible. We thus implicitly assume that W is continuous, although this restriction could probably be removed by using the concept of a generalized derivative, as in Brown (1968). Nonrandomized $\mathscr{G}$-equivariant estimators are of the form cS^m, and, assuming that $E_{0,1}W(cS^m)$ is not a monotone function of c, there exists an optimum choice of c, c^0, as established in the following lemma.

LEMMA 3.1. $E_{0,1}W(cS^m)$ *and* $E_{\lambda,1}[W(cS^m) \mid 0 < Z \leqq r]$ *are strictly bowl-shaped functions of* c, *assuming their respective minimum values at* c^0 *and* $c_\lambda(r)$ *satisfying*

$$E_{0,1}W'(c^0S^m)S^m = 0$$

and

$$E_{\lambda,1}[W'(c_\lambda(r)S^m)S^m \mid 0 < Z \leqq r] = 0 .$$

PROOF. The proof follows the lines of Lemma 2(iii), page 74 of Lehmann (1959), and uses the monotone likelihood ratio properties of $f_S(s/c \mid 0, 1)$ and $f_{S|0<Z\leqq r}(s/c \mid \lambda, 1)$. □

REMARK. For strictly convex loss the proof is immediate.

For fixed $r > 0$, and using monotone likelihood ratio properties of $\{f_{S|0<Z\leqq r}(s \mid \lambda, 1), f_S(s \mid 0, 1)\}$, it is not hard to see that

$$\sup_\lambda c_\lambda(r) = c_0(r) < c^0 , \tag{3.1}$$

and, as a consequence of Lemma 3.1, we have the following lemma.

LEMMA 3.2. *$\phi_r(Z)S^m$ has uniformly smaller risk than c^0S^m, where*

$$\begin{aligned} \phi_r(z) &= c_0(r) && 0 < z \leqq r \\ &= c^0 && z \leqq 0 \text{ or } z > r . \end{aligned} \tag{3.2}$$

Finally, using a monotone likelihood ratio argument again, we see that $c_0(r)$ is an increasing function of r. We therefore obtain the analogue of Theorem 2.1.

THEOREM 3.1. *The risk function of $\phi^{**}(Z)S^m$ is nowhere larger than that of c^0S^m, where*

$$\begin{aligned} \phi^{**}(z) &= c_0(z) && z > 0 \\ &= c^0 && z \leqq 0 . \end{aligned} \tag{3.3}$$

REMARK. The estimators considered in this paper are minimax (see, for example, Brewster (1972)).

4. Numerical comparisons of the risk functions. Figure 1 compares the risk

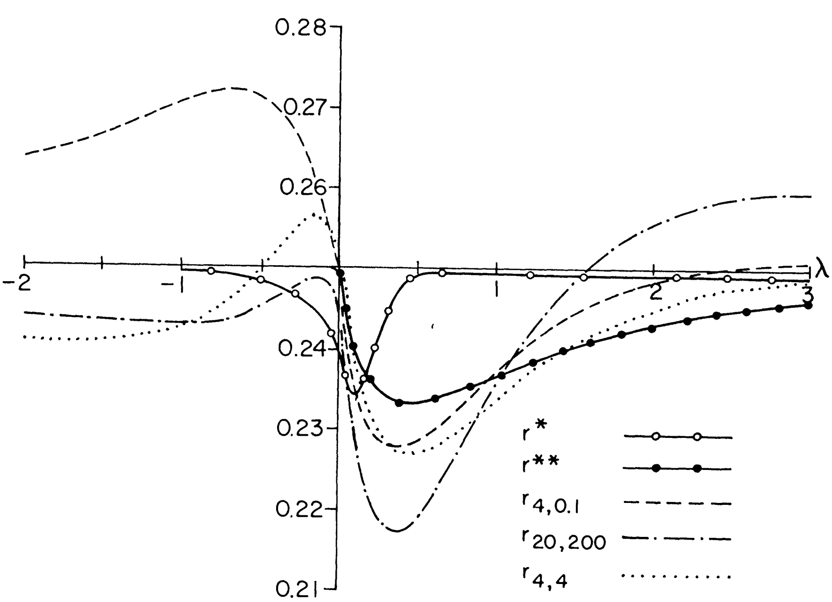

FIG. 1. A comparison of the risk functions when $n = 4$.

functions, $r^*(\lambda)$ and $r^{**}(\lambda)$, of $\phi^*(Z)S$ and $\phi^{**}(Z)S$, respectively. Here, loss is squared error and $n = 4$. The risk function of S is identically equal to .25.

For comparison, the risk functions of three estimators which are admissible within the class of $\mathscr{S}$-equivariant estimators are also presented. These estimators, $T_{a,b}(Z)S$, are generalized Bayes within the class of $\mathscr{S}$-equivariant estimators, with respect to priors on λ of the form

$$\begin{aligned}\pi_{a,b}(\lambda) &= (1 + a|\lambda|)^{-1} \qquad \lambda \geqq 0 \\ &= (1 + b|\lambda|)^{-1} \qquad \lambda < 0 .\end{aligned}$$

The scale-admissibility of these estimators was demonstrated by Zidek (1973). From Figure 1 it is apparent that these estimators are not minimax.

5. Acknowledgment. I am indebted to Professor J. V. Zidek for performing the numerical integration in Section 4, and for permitting me to include the risk functions of some of his scale-admissible estimators for comparison.

REFERENCES

[1] ARNOLD, B. C. (1970). Inadmissibility of the usual scale estimate for a shifted exponential distribution. *J. Amer. Statist. Assoc.* **65** 1260–1264.

[2] BREWSTER, J. F. and ZIDEK, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38.

[3] BREWSTER, J. F. (1972). Minimax estimation in location-scale problems. Technical Report No. 38, Dept. of Statistics, Univ. of Manitoba.

[4] BROWN, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.* **39** 29–48.

[5] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[6] SACKS, J. (1963). Generalized Bayes solutions in estimation problems. *Ann. Math. Statist.* **34** 751–768.

[7] STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

[8] ZIDEK, J. V. (1973). Estimating the scale parameter of the exponential distribution with unknown location. *Ann. Statist.* **1** 264–278.

DEPARTMENT OF STATISTICS
UNIVERSITY OF MANITOBA
WINNIPEG, MANITOBA, CANADA