ESTIMATION OF LINEAR PARAMETRIC FUNCTIONS FOR SEVERAL EXPONENTIAL SAMPLES

BY

A. L. Rukhin[1]
Purdue University

and

J. V. Zidek[2]
University of British Columbia

Technical Report #83-10

Revised August 1984

Purdue University
West Lafayette, Indiana

[1]Research supported by NSF Grant #MCS 78-02300.

[2]Research supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

ESTIMATION OF LINEAR PARAMETRIC FUNCTIONS FOR SEVERAL EXPONENTIAL SAMPLES

A. L. Rukhin[1], J. V. Zidek[2]

ABSTRACT. This work yields improved estimators for quadratic loss of linear functions, $\zeta = \sum_{j=1}^{m} \alpha_i \xi_i + \delta\sigma$, of parameters indexing m independent, exponential distributions. Here $\alpha_i \geq 0$, $\delta > 0$ are assumed known, while ξ_i and σ respectively, denote the unknown left hand endpoint and scale of the ith exponential population from which a sample of n_i independent observations is available. In the case when $\delta > (\Sigma n_i + 1)\ \Sigma\alpha_i/(mn_i)$ or $\zeta = \alpha_1\xi_1 + \delta\sigma$ and $0 \leq \delta < \alpha_1/n_1$, we construct an alternative to the best affine equivariant estimator, $\hat{\zeta}_0$, the risk of which is never more and actually less than that of $\hat{\zeta}_0$ over much of the parameter space.

1. Research supported by NSF Grant #MCS 78-02300.
2. Research supported by a grant from the Natural Sciences and Engineering Research Council of Canada and NSF Grant #MCS 83-01807.

AMS 1980 subject classifications. Primary 62C15; Secondary 62F10, 62H12, 62E15, 62C20.

Keywords and phrases. Exponential distribution, quantile estimation, location-scale parameters, inadmissibility, minimaxity, quadratic loss.

1. INTRODUCTION

This paper is concerned with the estimation of linear parametric functions of m independent two-parameter exponential populations. Samples of independent observables, $X_{i1},\ldots,X_{in_i}$, are available from the i-th of these, the population density function of which is

$$\exp\left[-(x-\xi_i)\sigma^{-1}\right]\sigma^{-1} \text{ or } 0 \text{ according as } x \geq \xi_i \text{ or } < \xi_i,\ i=1,\ldots,m.$$

Our work is aimed at finding a superior alternative to the best affine equivariant estimator of the parametric function

$$\zeta = \alpha\xi' + \delta\sigma = \Sigma\alpha_i\xi_i + \delta\sigma \tag{1.1}$$

of $\xi = (\xi_1,\ldots,\xi_m)$ and σ, for known α and δ. Of particular interest are the choices (i) $\alpha = e_j$, the jth basis vector, (ii) $\alpha = 0$ and (iii) $\alpha = (m^{-1},\ldots,m^{-1})$. In (i) ζ is a quantile of the j-th population when $\delta > 0$. In (ii) $\zeta = \sigma$, the scale parameter when $\delta = 1$. In (iii) one obtains $\bar{\xi} + \delta\sigma = m^{-1}\Sigma\xi_i + \delta\sigma$ which is of statistical interest. However, it is not possible to derive results which cover all possible choices of α without reducing the analysis to a succession of special cases. For brevity we restrict our analysis to the case of greatest statistical interest,

$$\alpha_i \geq 0,\ \delta \geq 0,\ i = 1,\ldots, m, \tag{1.2}$$

for which certain general results may be stated.

An estimator's performance is assumed to be measured by the quadratic loss function, $(\hat{\zeta} - \zeta)^2/\sigma^2$. With respect to this loss function, the best affine equivariant estimator is $\hat{\zeta}_0 = \alpha X' + d_0 Y$ (see Section 2) where

$$X = (X_1,\ldots,X_m),\quad X_i = \min\{X_{ij}: 1 \leq j \leq n_i\} \tag{1.3}$$

$$Y = (\Sigma\Sigma X_{ij} - \Sigma n_i X_i)/n,\ n = \Sigma n_i, \text{ and } d_0 = n(\delta - \Sigma\alpha_i/n_i)/(n-m+1).$$

Our objective is to construct an estimator, $\hat{\zeta}$, for which $E(\hat{\zeta}-\zeta)^2 \leq E(\hat{\zeta}_0-\zeta)^2$ for all ξ and σ with strict inequality for some parameter values. In other words we prove the inadmissibility of $\hat{\zeta}_0$ for some values of α and δ.

In the remainder of this section we state some implications of our findings. Related earlier work is acknowledged and described in Section 4. Other aspects of the problem are discussed there as well.

Let $\alpha = 0$ and $\delta = 1$ in equation (1.1), so that $\hat{\zeta}_0 = d_0Y = nY/(n-m+1)$. Then (see Corollary 1)

$$\hat{\zeta} = \begin{cases} d_0Y[1 - \gamma + \gamma \min\{1,T\}], & X_i > 0 \text{ for all } i \\ d_0Y, & \text{otherwise} \end{cases} \tag{1.4}$$

where $T = [\Sigma\Sigma X_{ij}/(n+1)]/[(\Sigma\Sigma X_{ij} - \Sigma n_iX_i)/(n-m+1)]$ is better than $\hat{\zeta}_0$ if $0 < \gamma \leq 2$. Observe that T^{-1} is, essentially, the likelihood ratio for testing H_0: $\xi_1 = \ldots = \xi_m = 0$. When $\gamma = 1$ our procedure, $\hat{\zeta}$, takes values, $\Sigma\Sigma X_{ij}/(n+1)$ or $(\Sigma\Sigma X_{ij} - \Sigma n_iX_i)/(n-m+1)$ according as this preliminary test accepts or rejects H_0. Stein (1964) gave an analogous interpretation of his estimator of the normal variance.

It would seem plausible that whenever δ is "large" relative to α, similar improvements should be possible. Indeed they are found in Corollary 1. If $\delta > (n+1)\Sigma\alpha_i/(mn_i)$ then

$$\hat{\zeta} = \begin{cases} \alpha X' + d_0Y[1 - \gamma + \gamma \min\{1,\delta T/(\delta-\Sigma\alpha_i/n_i)\}], & \text{if } X_i > 0 \text{ or all } i \\ \alpha X' + d_0Y, & \text{otherwise} \end{cases} \tag{1.5}$$

is better than $\hat{\zeta}_0$ where T is defined in equation (1.4) and $0 < \gamma \leq 2$. As in equation (1.4) the significance level, α^*, of the preliminary test in equation (1.5) does not depend on γ, $0 < \gamma \leq 2$. However, in the latter case it does depend on $\Sigma\alpha_i/(\delta n_i)$. As δ decreases to $(n+1)\Sigma\alpha_i/(mn_i)$, the acceptance region contracts from that above for $\alpha = 0$, $\{T < 1\}$, to $\{T < 1 - m/(n+1)\}$ which maximizes α^*. Under H_0, statistic T has the representation $T = [1-m/(n+1)]\ F$ where F has an F distribution and $\alpha^* = P[F < 1]$.

These results show that gains in estimating ζ when δ is large are due to better estimating its component σ. The size of the preliminary test's acceptance region (the region where improvements over $\hat{\zeta}_0$ are possible) is a decreasing function of $\Sigma\alpha_i/(\delta n_i) < 1$. Improvements are more difficult to achieve when this latter quantity is large than when it is small.

If δ is a small positive number, our methods fail unless α is a scalar, say $\alpha\xi' = \alpha_1 \xi_1$. The alternatives to $\hat{\zeta}_0$ which are given in Corollaries 2 and 3 do not seem to admit any simple interpretation. The gains which they achieve are not made, in particular, by implicitly tuning the estimator of σ.

2. TECHNICAL PRELIMINARIES

The sufficiency principle reduces the problem to observables (X,Y) where X and Y are defined in equation (1.3). The joint density function of (X,Y) is $\sigma^{-m-1}p(u,v)$ where $u = (x-\xi)/\sigma$, $v = y/\sigma$ and

$$p(u,v) \propto v^{n-m-1} \exp(-\Sigma n_i u_i - nv),\ n > m.$$

A fully equivariant estimator, $\hat{\zeta}$, of $\zeta = \alpha\xi' + \delta\sigma$, $\alpha \geq 0$, transforms as $\hat{\zeta} \to a\hat{\zeta} + \alpha c$ when $X \to aX + c$ and $Y \to aY$ $(a > 0)$ since the parameters transform as $\xi \to a\xi + c$ and $\sigma \to a\sigma$ while $\zeta \to a\zeta + \alpha c$. It is easily shown that any of these procedures has the form $\hat{\zeta} = \alpha X' + dY$, for some real d. Moreover, the examination of the risk of these estimators shows that the best choice within this class is

$$\hat{\zeta}_0 = \alpha X' + d_0 Y, \tag{2.1}$$

where $d_0 = n(n-m+1)^{-1}(\delta - \Sigma\alpha_j n_j^{-1})$.

We study the inadmissibility of a more general scale-equivariant estimator, $\hat{\zeta}^* = \phi_0(Z)Y$, $Z = X/Y$. Notice that in the case of the best equivariant estimator,

$$\phi_0(Z) = \alpha Z' + d_0.$$

Guided by the concluding remarks of Stein (1964) we look for an improvement upon $\hat{\zeta}^*$ of the form

$$\hat{\zeta} = \hat{\zeta}^* - \phi(Z)Y$$

where ϕ is a real-valued measurable function. The risk of this procedure depends on (ξ,σ) only through ξ/σ so it may be assumed without loss of generality that $\sigma = 1$.

The proof of the following Lemma is based on the inequality $x(x-2) < 0$ for $0 < x < 2$, and we omit it. Notice however that in our situation this result is more precise than Theorem 3.3.1 of Brewster and Zidek (1974) which deals with arbitrary bowl-shaped loss functions.

LEMMA 1. Suppose that there exists a measurable function $W = W(Z)$ such that ϕ_0 is a function of W and

$$2 \sup_{\xi} d(\xi,W) \leq \phi(W) \leq 0$$

or

$$0 \leq \phi(W) \leq 2 \inf_{\xi} D(\xi,W)$$

where

$$D(\xi,W) = \phi_0 - (\alpha\xi' + \delta)T$$

and

$$T = T(\xi,W) = E_\xi(Y|W)/E_\xi(Y^2|W).$$

If $\hat{\zeta} = Y\phi_0 - Y\phi(W)$ is different from $\hat{\zeta}^*$, with positive probability, then it improves upon $\hat{\zeta}^*$.

In the sequel, W is typically Z or one of the coordinates of Z. More generally let $W = (Z_1,\ldots,Z_\ell)^T$, $\ell \leq m$. Define $J(+) = \{j\colon 1 \leq j \leq \ell,\ Z_j > 0\}$, $J(-) = \{j\colon 1 \leq j \leq \ell,\ Z_j < 0\}$ and let $A = \max\{0, \max\{\xi_j/Z_j\colon j \in J(+)\}\}$ if $J(+)$ is nonempty, $A = 0$ otherwise. Analogously let $B = \min\{\xi_j/Z_j\colon j \in J(-)\}$ if $J(-)$ is nonempty and $B = \infty$ otherwise. Then $A < B$ with probability 1 and for given $W = w$, T depends on ξ and w only through A and B since

$$T = \int_I y f_w(y)dy \Big/ \int_I y^2 f_w(y)dy, \tag{2.2}$$

where $I = (A,B)$, $f_w(y) = y^{N-1}\exp(-ys)$, $N = n+\ell-m$ and $s = n + \sum_{j=1}^{\ell} n_j z_j = s(w)$.

It is easily seen that for a fixed $W = w$ and with an abuse of notation, $T = T(A,B)$ is decreasing in A and B. We now consider the special case, $A = 0$ and $s \leq 0$. Let K denote a Poisson random variable with mean $t = |s|B$.

LEMMA 2. If $A = 0$ and $s \leq 0$,

$$T = (|s|/t)E(N+K+1)^{-1}/E(N+K+2)^{-1}. \tag{2.3}$$

PROOF. Since $A = 0$ and $s \leq 0$, $T = |s| \int_0^t y^N \exp(y)dy / \int_0^t y^{N+1} \exp(y)dy$.

But

$$\int_0^t y^N e^y dy = \exp(t)t^{N+1} \int_0^1 u^N \exp[-t(1-u)]du,$$

and $e^{-t(1-u)} = Eu^K$. The remainder of the argument is now obvious.

The proof of the next result is straightforward.

LEMMA 3. For any function h such that $E|h(K)| < \infty$,

$$tEh(K) = EKh(K-1). \tag{2.4}$$

These results now lead to

LEMMA 4. If $s \leq 0$, $c \leq 0$, and $A = 0$,

$$a - (aB+c)T \geq -a(N + \tilde{K} + 1)^{-1} + cs(\tilde{K} + 1)^{-1} \tag{2.5}$$

where

$$\tilde{K} = \begin{cases} 0, & \text{if } csA_1 \leq a \\ r, & \text{if } csA_{r+1} \leq a < csA_r \end{cases} \tag{2.6}$$

and $A_r = (N+r)(N+r+1)(r)^{-1}(r+1)^{-1}$, $r = 1,2,\ldots$.

PROOF. By Lemma 2, $(aB+c)T = (at+c|s|)t^{-1} \times E(N+K+1)^{-1}/E(N+K+2)^{-1}$. Lemma 3 implies that $tE(N+K+2)^{-1} = EK(N+K+1)^{-1}$ and $tE(N+K+1)^{-1} = EK(N+K)^{-1}$. So $a - (aB+c)T = a - EK(N+K+1)^{-1}[a+a(N+K)^{-1} - csK^{-1}]/EK(N+K+1)^{-1} = -E[a/(N+K^*) - cs/K^*]$ where $P[K^* = k] \propto kP(K=k) / (N+k+1)$. Thus $a - (aB+c)T \geq -\max\{H(k): k \geq 1\}$ where $H(x) = a/(N+x) - cs/x$, $x \geq 0$. Since H has a unique maximum, the conclusion follows.

<u>LEMMA 5.</u> <u>If</u> $z_j > 0$ for all j, $W = z$ <u>and</u> $\delta > 0$

$$\inf_{\xi} D(\xi,W) \geq \phi_0 - \alpha z' - \delta s(N+1)^{-1}.$$

PROOF. Observe that the infimum in the condition of Lemma 1 is

$$\phi_0 - \sup(\alpha\xi' + \delta)T \geq \phi_0 - \sup \alpha\xi' T - \sup \delta T.$$

But $\alpha\xi' T \leq \alpha z' \max (\xi_j/z_j) T \leq \alpha z' AT(A,\infty) \leq \alpha z'$ since $AT(A,\infty) \leq 1$ and this with $T \leq s(N+1)^{-1}$ implies the conclusion.

The last technical result which we shall need deals with the case $\delta > 0$, $\alpha = \alpha_1 > 0$ (a scalar) and $W = Z_1$. In this case $s = n + n_1 z_1$.

<u>LEMMA 6.</u> <u>If</u> $\delta > 0$ <u>and</u> $Z = z$ <u>is given with</u> $s = n+n_1 z_1 \leq 0$, <u>then</u>

$$\sup_{\xi} D(\xi,W) \leq \phi_0 - \alpha_1 z_1 (N+2)/(N+1) \tag{2.7}$$

and

$$\sup D(\xi,W) \leq \phi_0 - \alpha_1 z_1 (N + \tilde{K} + 2)/(N + \tilde{K} + 1) + \delta s/(\tilde{K}+1), \tag{2.8}$$

<u>where</u> $\tilde{K}$ <u>is given in equation</u> (2.6) <u>with</u> $a = \alpha_1 |z_1|$, $c = -\delta$ <u>and</u> $N=n-m+1$.

PROOF. This is a straightforward application of the easily derived inequality $BT(0,B) \leq (N+2)(N+1)^{-1}$, and Lemma 4.

3. PRINCIPAL RESULTS

The first of these pertains to the case when δ is large.

THEOREM 1. *Let* $s = n + \Sigma n_i z_i$ *and for some* γ, $0 \le \gamma \le 2$,

$$\hat{\zeta} = Y[\phi_0 - \gamma \max\{0, \phi_0 - \alpha z' - \delta s(N+1)^{-1}\}] \tag{3.1}$$

if $z_j > 0$ for all j, *and* $\hat{\zeta} = \hat{\zeta}^* = \phi_0 Y$ *otherwise*. Then $\hat{\zeta}$ is better than $\hat{\zeta}^*$ *if* $\hat{\zeta} \neq \hat{\zeta}^*$ *on a set of positive probability*.

PROOF. Apply Lemmas 1 and 5.

As an application of this result, consider the case $\hat{\zeta}^* = \hat{\zeta}_0$, i.e., $\phi_0 = \alpha z' + d_0$. Then $\hat{\zeta}$ of equation (3.1) is different from $\hat{\zeta}_0$ if for a set of z's of positive probability, $d_0 - \delta s(N+1)^{-1} > 0$, i.e., if $\delta > (N+1)m^{-1}\Sigma\alpha_j/n_j$. This yields the following result.

COROLLARY 1. *If* $\delta > (N+1)m^{-1}\Sigma\alpha_j/n_j$ *then, given* $Z = z$,

$$\hat{\zeta} = \hat{\zeta}_0 - \gamma Y \max\{0, d_0 - \delta s(N+1)^{-1}\} \tag{3.2}$$

is better than $\hat{\zeta}$ *if* $0 < \gamma \le 2$.

In the next theorem we construct improvements of the kind given in Theorem 1.

THEOREM 2. *Suppose* $\delta > 0$ *and* $Z = z$ *is given*. *Let*

$$\hat{\zeta} = Y[\phi_0 - \gamma \min\{0, \bar{D}\}]$$

if $z_1 < -n/n_1$, $0 \le \gamma \le 2$ and $\hat{\zeta} = \hat{\zeta}^* = \phi_0 Y$ *otherwise where* $\bar{D}$ *is either of the upper bounds given in equations* (2.7) *and* (2.8). *Then* $\hat{\zeta}$ *is better than* $\hat{\zeta}^*$ *if it is different from* $\hat{\zeta}^*$, *with positive probability*.

PROOF. This is an immediate application of Lemmas 1 and 6.

Theorem 2 may now be applied to the case of prinicpal interest, $\hat{\zeta} = \hat{\zeta}_0$.

COROLLARY 2. *If* $0 \le \delta < \alpha_1 n_1^{-1}(n-m+2)^{-1}$, $Z = z$, *then*

$$\hat{\xi} = \begin{cases} \hat{\zeta}_0 - \gamma\, Y \min\{0,\, d_0 - \alpha_1 z_1/(N+1)\}, & z_1 \le -n/n_1 \\ \hat{\zeta}_0 & \text{otherwise} \end{cases}$$

is better than $\hat{\zeta}_0$ *when* $0 < \gamma \le 2$, *where* $N = n-m+1$.

PROOF. Apply Theorem 2. Under the hypothesized constraints on δ, $\hat{\zeta} \ne \hat{\zeta}_0$ on a set of positive probability when $\xi_1 < 0$.

To simplify the statement of the last result in this Section let

$$\Delta = \delta n_1/\alpha_1 \text{ and } f_r = r(N+r)^{-1},\ r = 0,1,2,\ldots \text{ where } N=n-m+1.$$

COROLLARY 3. *Assume* $\Delta < 1$ *and that* $\Delta \ne f_r$ *for any* r. *Let*

$$\hat{\zeta} = \hat{\zeta}_0 - \gamma\, Y \min\{0,\, d_0 - \alpha_1 z_1/(N+r+1) + \delta(n+n_1 z_1)/(r+1)\}$$

when

$$-(\Delta - f_{r+1} f_{r+2}) z_1 \le n\Delta n_1^{-1} < -(\Delta - f_r f_{r+1}) z_1$$

and $\hat{\zeta} = \hat{\zeta}_0$ *otherwise*.

Then $\hat{\zeta}$ *is better than* $\hat{\zeta}_0$ *for* $0 < \gamma \le 2$.

PROOF. The proof consists of showing that $\hat{\zeta}$ is different from $\hat{\zeta}_0$ with positive probability when the bound given in (2.8) is used. In the present case, $\phi_0 = \alpha_1 Z_1 + d_0$ where $d_0 = nN^{-1}(\delta - \alpha_1 n_1^{-1})$ and $N = (n-m+1)$. Thus $d_0 = n\alpha_1 n_1^{-1} N^{-1}(\Delta-1)$ and the upper bound in (2.8) becomes $d_0 - \alpha_1 z_1 (N+\tilde{K}+1)^{-1} + \delta(n+n_1 z_1)(\tilde{K}+1)^{-1} = n\alpha_1 n_1^{-1}[(\Delta-1)N^{-1} + u((N+\tilde{K}+1)^{-1} - \Delta(\tilde{K}+1)^{-1}) + \Delta(\tilde{K}+1)^{-1}]$ where $\tilde{K} = 0$ or r according as $(\Delta - f_1 f_2)u \le \Delta$ or $(\Delta - f_{r+1} f_{r+2})u \le \Delta < (\Delta - f_r f_{r+1})u$, $r = 1, 2, \ldots$, $u = -n_1 Z_1/k$.

Let $R = \min\{r \ge 1: \Delta - f_r f_{r+1} \le 0\}$. Then $0 \le \tilde{K} < R$ if $u > 1$, probability of which is positive. The proof reduces to showing that with positive probability, $(\Delta - 1)N^{-1} + u[(N+\tilde{K}+1)^{-1} - \Delta(\tilde{K}+1)^{-1}] + \Delta(\tilde{K}+1)^{-1} = B$, say, is negative for some $\tilde{K} = r < R$ and an associated set of u-values which has positive probability for each of a set of parameter values.

If $R = 1$, i.e. $0 < \Delta \le f_1 f_2$, then $\tilde{K} = 0$ with probability 1. In this case, $B = (\Delta-1)N^{-1} + u[(N+1)^{-1} - \Delta] + \Delta$ and, since $\Delta \le f_1 f_2 < f_1 = (N+1)^{-1}$, B's infimum for $u \ge 1$ is attained at $u = 1$ where $B = N^{-1}[\Delta - (N+1)^{-1}] < 0$. So $\hat{\zeta}$ differs from $\hat{\zeta}_0$ with positive probability and the proof is complete.

Now suppose $f_K f_{K+1} < \Delta \le f_{K+1} f_{K+2}$ so that $R = K+1$, $1 \le K < \infty$. Then $\tilde{K} = r$ when $L_r < u \le L_{r+1}$ where $L_r = \Delta(\Delta - f_r f_{r+1})^{-1}$, $0 \le r \le K$ and $L_{K+1} = \infty$. In this case, $B = (\Delta-1)N^{-1} + u[(N+r+1)^{-1} - \Delta(r+1)^{-1}] + \Delta(r+1)^{-1}$ and this is minimized at either $u = L_r$ or L_{r+1} for $0 \le r < K$. At $r = K$, it is minimized at $u = L_K$ since $(N+K+1)^{-1} - \Delta(K+1)^{-1} > 0$ and $L_K \le u$ but this fact will not be needed. At $u = L_r$, $B = (\Delta-1)N^{-1} + H(r)$ while at $u = L_{r+1}$, $r < K$, $B = (\Delta-1)N^{-1} + H(r+1)$ where $H(r) = N\Delta[\Delta(N+r)(N+r+1) - r(r+1)]^{-1}$, for all r.
For a continuous argument, x, $H(x)$ is minimized when $x = \zeta = [N\Delta - \tfrac{1}{2}(1-\Delta)][1-\Delta]^{-1}$, i.e., $(\zeta + \tfrac{1}{2})(N+\zeta+\tfrac{1}{2})^{-1} = \Delta$. But $\zeta < K$ since otherwise, $K(N+K)^{-1} < (\zeta+\tfrac{1}{2})(N+\zeta+\tfrac{1}{2})^{-1} \le (K+1)(K+2)(N+K+1)^{-1}(N+K+2)^{-1}$ which is impossible if $K(N+K) \ge 2$ as is true here where $K \ge 1$ and $N \ge 1$. So $r \le \zeta < r+1$ for some r, $o \le r < K$; let ρ denote

this value of r and $\varepsilon = \rho - \zeta$ so that $\rho = \zeta + \varepsilon$. Then $H(\rho) = (1-\Delta)N^{-1}\nu(\nu+\frac{1}{4}-\varepsilon^2)^{-1}$ $\nu = \Delta N^2(1-\Delta)^{-2}$, and $H(1+\rho) = (1-\Delta)N^{-1}\nu(\nu+\frac{1}{4}-[1+\varepsilon]^2)^{-1}$. If $H(\rho)<(1-\Delta)N^{-1}$, $B<0$ at $u=L_r$, $r = \rho$ and the proof is complete. Otherwise, $H(\rho) \geq (1-\Delta)N^{-1}$, i.e., $-1 < \varepsilon \leq -\frac{1}{2}$ so that $0 < 1 + \varepsilon \leq \frac{1}{2}$, $H(1+\rho) \leq (1-\Delta)N^{-1}$ and $B \leq 0$ at $r = 1 + \rho$ and $u = L_{1+\rho}$. In fact, $B < 0$ since $B = 0$ entails $\varepsilon = -\frac{1}{2}$, i.e. $\zeta = \rho + \frac{1}{2}$ or $\Delta = f_{\rho+1}$, a possibility which is ruled out by our hypothesis. So again the proof is complete.

This proof reveals why $\Delta < 1$ is required in the application of (2.8). For if $1 \leq \Delta$, B is minimized by $u = L_{r+1}$ when $K = r$. But then $B = (\Delta-1)N^{-1} + H(r+1) > 0$ for every r.

4. DISCUSSION.

The approach embodied in Lemma 1 for finding superior alternatives to best equivariant estimators and variants of it have been used elsewhere (see Stein 1964, Zidek 1971, Zidek 1973, Brewster 1974, Brewster and Zidek 1974, Shorrock and Zidek 1976, Sharma 1977, Tsui, Weerahandi and Zidek 1980 and Rukhin and Strawderman 1982). Brewster and Zidek (1974) point out that this method differs from those which yield estimators like the celebrated James-Stein rule. It seeks improvement on an orbit-by-orbit basis, whereas the superiority of the James-Stein estimators is achieved by integrating across orbits.

Our results for small $\delta \geq 0$ generalize those of Rukhin and Strawderman (1982) (hereafter RS), for the same case and our proof is a straight-forward adaption of theirs. They choose $\gamma = 2$ in this and the other cases, but this choice is somewhat arbitrary since there is no uniformly best choice among the rules indexed by γ.

The result of Corollary 1 when specialized to the case of quantile estimation and $m = 1$ differs from that of RS and is according to their numerical comparisons, inferior to it. Their analysis for large δ is a variant of ours, which judiciously incorporates the "integration by parts" method of Stein (1973), something we are unable to do when $m > 1$ for technical reasons. In effect, they transform ξ_1, X_1, Y into S, T which are independently distributed, T is ancillary (given $\sigma = 1$) and S has the gamma distribution with scale parameter $-\xi_1$ when $X_1 < 0$ and hence $\xi_1 < 0$. By this formal device the problem is forced into a form which is amenable to the application of Stein's method.

We have ignored the cases, $\delta = \alpha_1 r/[n-m+1+r]$ $r = 1,2,...$ As is observed in RS a superior alternative to $\hat{\xi}_0$ exists here as well.

It is not known if, when $\alpha_1/n_1 \leq \delta \leq (n+1)\alpha_1/(mn_1)$, the best equivariant estimator of $\alpha_1\xi_1 + \delta\sigma$ is admissible. Some support in favour of this proposition is given by RS when m=1 and Rukhin (1983) proves this result in the same case. On the other hand, even for δ's in this range it seems unlikely that the simultaneous estimator of $\alpha_i\xi_i + \delta_i\sigma$, $i = 1,\ldots, m$ would be admissible at least if $m \geq 3$ by analogy with the James-Stein result. This issue too remains open.

Here, as in all similar studies, it is unclear whether the gains from fine-tuning the parameter estimates are worthwhile when measured against the inevitable approximation error introduced by the choice of the underlying sampling model. So the potential practical impact of such work remains unclear, even though the exponential model is often assumed in the reliability theory (for references see RS). Such work would seem to be endowed with theoretical interest because it belongs to the theory of multiparameter estimation. Results such as that of James and Stein (1961) indicate that large sample theory may well be misleading when applied to moderate samples from multi-dimensional statistical models. The two-parameter exponential model is of particular interest in this context, because of its structural irregularities, combined with its mathematical tractability.

ACKNOWLEDGEMENT. We are indebted to a referee for an extremely careful review.

REFERENCES

Brewster, J.F. Alternative estimators for the scale parameter of the exponential distribution with unknown location. Ann. Statist., 2 553-557 (1974).

Brewster, J.F. and Zidek, J.V. Improving on equivariant estimators. Ann. Statist., 2, 21-38 (1974).

James, W. and Stein, C. Estimation with quadratic loss. Proc. Fourth Berk. Symp. Math. Statist. Probability, 1, 361-379. Univ. of California Press (1961).

Rukhin, A.L. and Strawderman, W.F. Estimating a quantile of an exponential distribution. Jour. Amer. Statist. Assoc., 77, 159-162 (1982).

Rukhin, A.L. Admissibility and minimaxity results in the estimation problem of exponential quantiles. Tech. Rep. #83-54, Department of Statistics, Purdue University (1983).

Sharma, D. Estimation of the reciprocal of the scale parameter in a shifted exponential distribution, Sankhya, Ser. A, 39, 203-205 (1977).

Shorrock, R.W. and Zidek, J.V. An improved estimator of the generalized variance. Ann. Statist., 4, 629-638 (1976).

Stein, C. Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Instit. Statist. Math., 16, 155-160 (1964).

Stein, C. Estimation of the mean of a multivariate normal distribution. Proc. Prague Symp. Asymptotic Statist., 2, 345-381 (1973).

Tsui, K-W., Weerahandi, S. and Zidek, J.V. Inadmissibility of the best fully equivariant estimator of the generalized residual variance. Ann. Statist., 8, 1156-1159 (1980).

Zidek, J.V. Inadmissibility of a class of estimators of a normal quantile. Ann. Math. Statist., 42, 1444-1447 (1971).

Zidek, J.V. Estimating the scale parameter of the exponential distribution with unknown location. Ann. Statist., 1, 264-278 (1973).

A. L. Rukhin
Department of Statistics
Purdue University
West Lafayette, Indiana 47907 USA

J. V. Zidek
Department of Statistics GN-22
University of Washington
Seattle, Washington 98195