The Annals of Mathematical Statistics
1968, Vol. 39, No. 2, 502–516

ESTIMATION OF THE LARGER TRANSLATION PARAMETER

By Saul Blumenthal[1] and Arthur Cohen[2]

New York University and Rutgers-The State University

1. Introduction and summary. Let the random variables $X_{i1}, X_{i2}, \cdots, X_{in}$, $i = 1, 2$, be real valued and independent with density functions $f(x - \theta_i)$ (θ_i real), $i = 1, 2$, (with respect to Lebesgue measure). We take $\int_{-\infty}^{\infty} xf(x)\,dx = 0$ with no loss of generality. The problem considered here is estimation of the function $\varphi(\theta_1, \theta_2) = \text{maximum}\,(\theta_1, \theta_2)$ with a squared error loss function. Questions of minimaxity and admissibility of certain natural estimators are considered. This problem is the estimation analogue of the well known ranking and selection problem, which has received considerable attention in the past. For a bibliography see Bechhofer, Kiefer, and Sobel (1968). Whereas previous work is concerned with choosing the population with the larger parameter, here we are concerned with estimating the larger parameter.

Consider for the moment, the case where $n = 1$. A natural estimator of $\varphi(\theta_1, \theta_2)$ is $\varphi(X_{11}, X_{21}) = \text{maximum of } (X_{11}, X_{21})$. This estimator is symmetric in the observations and invariant under translations which take (θ_1, θ_2) to $(\theta_1 + a, \theta_2 + a)$, for any real constant a. Under suitable conditions, one of which is that f be symmetric, (the conditions will be stated precisely for general n in the next paragraph), it is shown that $\varphi(X_{11}, X_{21})$ is minimax. However $\varphi(X_{11}, X_{21})$ is not in general admissible. Furthermore, when f is not symmetric $\varphi(X_{11}, X_{21})$ also need not be minimax. Another estimate with the invariance properties stated above is the *a posteriori* expected value of $\varphi(\theta_1, \theta_2)$, given (X_{11}, X_{21}), when the generalized prior distribution of (θ_1, θ_2) is taken to be the uniform distribution over the two dimensional plane. For many estimation problems, such an adaptation of the Pitman estimator is known to be minimax. However, for this problem, under suitable conditions (again including that f be symmetric), this estimator need not be minimax. It is true though, that this estimator is admissible.

In order to summarize the results for general n, it is convenient to define the analogues of the estimates considered above. Let X_i be the *a posteriori* expected value of θ_i, given X_{ij}, $j = 1, 2, \cdots, n$, when θ_i has the generalized uniform distribution as a prior distribution. That is,

$$X_i = \int \theta_i \textstyle\prod_{j=1}^n f(X_{ij} - \theta_i)\,d\theta_i \Big/ \int \prod_{j=1}^n f(X_{ij} - \theta_i)\,d\theta_i, \quad i = 1, 2. \tag{1.1}$$

Note X_i is the usual Pitman estimator of θ_i and thus if f is the normal density

Received 18 January 1967; revised 13 November 1967.

[1] Research supported in part by NSF Grant GP 4933 while the author was at Rutgers-The State University, and in part by NSF Grant GP 7024 at New York University.

[2] Research supported by NSF Grant GP 4644.

then $X_i = \bar{X}_i$. Let

$$(1.2)\quad \varphi(X_{11}, X_{12}, \cdots, X_{1n}, X_{21}, \cdots, X_{2n}) = \varphi(X_1, X_2) = \text{maximum of } (X_1, X_2).$$

Also define δ^* to be the a *posteriori* expected value of $\varphi(\theta_1, \theta_2)$, given X_{ij}, $i = 1, 2, j = 1, 2, \cdots, n$, when the generalized prior distribution of (θ_1, θ_2) is the uniform distribution on the two dimensional plane. That is,

$$(1.3)\quad \begin{aligned} \delta^*(X_{11}, X_{12}, \cdots, X_{2n}) &= \textstyle\iint \varphi(\theta_1, \theta_2) \prod_{j=1}^n f(X_{1j} - \theta_1) \prod_{j=1}^n f(X_{2j} - \theta_2)\, d\theta_1\, d\theta_2 \\ &\quad \cdot \textstyle[\iint \prod_{j=1}^n f(X_{1j} - \theta_1) \prod_{j=1}^n f(X_{2j} - \theta_2)\, d\theta_1\, d\theta_2]^{-1}. \end{aligned}$$

In order to explicitly state the main results it is convenient to introduce some notation which will be more formally presented in the next section. Recall X_i is the usual Pitman estimator of θ_i. Let $Y_i = (Y_{i1}, Y_{i2}, \cdots, Y_{i,n-1})$ where $Y_{ij} = X_{i,j+1} - X_{i1}$. Let $p(x, y)$ be the conditional density of X_i given Y_i when $\theta_i = 0$.

Now for the problem of estimating $\varphi(\theta_1, \theta_2)$ the main results are as follows: (a) If $p(x, y) = p(-x, y)$ and $EE[(X_1^2 + X_2^2) \mid Y_1, Y_2] < \infty$, then $\varphi(X_1, X_2)$ is minimax. The proof of this result is based on an idea of Farrell (1964). (b) If $p(x, y) \neq p(-x, y)$, then the problem is complicated and in fact an example is given showing that $\varphi(\cdot)$ is not minimax. (c) The estimator $\varphi(\cdot)$ need not be admissible. When f is normal, it is easily shown, using a theorem and remark of Sacks (1963) that $\varphi(\cdot)$ is inadmissible (Section 4). (d) if $E\{E^2[(X_1^2 + X_2^2) \cdot |\log (X_1^2 + X_2^2)|^\beta \mid Y_1, Y_2]\} < \infty$ for some $\beta > 0$, then $\delta^*(\cdot)$ is admissible. The proof of this fact depends on results of Stein (1959b) and James and Stein (1961). (e) The estimator δ^* need not be minimax and in Section 3 it is shown that frequently δ^* is not minimax.

Despite the inadmissibility of $\varphi(X_1, X_2)$, this estimator has great intuitive appeal, is easy to use, and we have not succeeded in finding an estimate whose risk improves on the risk of φ. Also it is minimax under the restrictions noted above. Furthermore, if the problem was to simultaneously decide which population has the larger mean and estimate the larger mean, then a reasonable formulation for which $\varphi(X_1, X_2)$ is an admissible estimator of $\varphi(\theta_1, \theta_2)$ could easily be found. See for example Cohen (1965). The admissible δ^*, on the other hand, is more difficult to compute and has a larger bias than $\varphi(X_1, X_2)$.

In a future paper, we shall discuss topics such as the general question of unbiased estimation of $\varphi(\theta_1, \theta_2)$, the maximum likelihood estimate and its properties, and the generation of invariant Bayes estimates of which δ^* is a particular example.

As we develop the main results, generalizations will be indicated. In the next section we introduce the notation for a more general model than as given earlier and develop some preliminaries. In Section 3 we give the minimax results while in Section 4 we give the admissibility results.

2. Notation. In this section, we introduce the notation to be followed hereafter and relate it to that of the Introduction.

Following Stein (1959a) we define the random variables X_i, Y_{ij} ($i = 1, 2$; $j = 1, \cdots, n-1$) by

$$X_i = X_{i1} - r_1(Y_{i1}, \cdots, Y_{i,n-1}), \qquad i = 1, 2, \tag{2.1}$$

and

$$Y_{ij} = X_{i,j+1} - X_{i1}, \qquad i = 1, 2;\ j = 1, \cdots, n-1, \tag{2.2}$$

where

$$r_1(y) = (r_0(y))^{-1} \int x f(x) f(x + y_1) \cdots f(x + y_{n-1})\, dx, \tag{2.3}$$

$$r_0(y) = \int f(x) f(x + y_1) \cdots f(x + y_{n-1})\, dx, \tag{2.4}$$

and

$$y = (y_1, \cdots, y_{n-1}). \tag{2.5}$$

The conditional density of X_i given $(Y_{i1}, \cdots, Y_{i,n-1})$ when $f(x)$ is the density of X_{ij} is

$$p(x, y) = \{f(x + r_1(y)) f(x + r_1(y) + y_1) \cdots f(x + r_1(y) + y_{n-1})\}/r_0(y) \tag{2.6}$$

and when $f(x - \theta_i)$ is the density of X_{ij}, the conditional density given the Y's is $p(x - \theta_i, y)$. It will be noted that $X_i = X_{i1}$ if $n = 1$, and is the Pitman estimator of θ_i (given by (1.1)) in general.

Hereafter, $\mathcal{Y}_i$ ($i = 1, 2$) will be arbitrary but identical spaces, $\nu(\cdot)$ will be a probability measure defined on the Borel subsets of $\mathcal{Y}_i$ ($i = 1, 2$), and x will be a real number. $p(\cdot, \cdot) \geqq 0$ on $E_1 \times \mathcal{Y}_i$ ($i = 1, 2$) to E_1 is jointly measurable in the two variables and

$$\int p(x, y)\, dx = 1 \quad \text{for all} \quad y \in \mathcal{Y}_i, \qquad (i = 1, 2), \tag{2.7}$$

$$\int x p(x, y)\, dx = 0 \quad \text{for all} \quad y \in \mathcal{Y}_i, \qquad (i = 1, 2). \tag{2.8}$$

(Note that p given by (2.6) satisfies (2.7) and (2.8) with $\mathcal{Y}_i = E_{n-1}$ and $\nu(dy)$ given by $r_0(y)\, dy_1 \cdots dy_{n-1}$.) We shall sometimes find it convenient to write $\mathcal{Y}$ for $\mathcal{Y}_1 \times \mathcal{Y}_2$, y for (y_1, y_2), ν for $\nu \times \nu$, and $p(x_1, x_2, y)$ for $p(x_1, y_1)p(x_2, y_2)$.

Hereafter, the observed variables are (X_i, Y_i) with $X_i \in E_i$, $Y_i \in \mathcal{Y}_i$ ($i = 1, 2$) and the marginal conditional density of X_i given Y_i is $p(x_i - \theta_i, y_i)(-\infty < \theta_i < \infty)$ ($i = 1, 2$).

The risk of an estimator $\delta(X_1, X_2, Y)$ for $\varphi(\theta_1, \theta_2)$ is given by

$$R(\theta_1, \theta_2, \delta) = \int_{\mathcal{Y}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\delta(x_1, x_2, y) - \varphi(\theta_1, \theta_2)]^2 \cdot p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 \nu(dy). \tag{2.9}$$

In terms of this notation the estimator (1.2) remains as

$$\varphi(X_1, X_2) = \max(X_1, X_2) \tag{2.10}$$

while the estimator (1.3) becomes

$$(2.11)\quad \delta^*(X_1, X_2, Y_1, Y_2) = \iint \varphi(\theta_1, \theta_2)p(X_1 - \theta_1, Y_1)p(X_2 - \theta_2, Y_2)\, d\theta_1\, d\theta_2 .$$

For the purposes of the next section, it will be convenient to use the following change of variables and corresponding change of parameters. Namely, let

$$(2.12a)\qquad Z_0 = (X_2 - X_1)/2, \qquad Z_1 = (X_1 + X_2)/2;$$

$$(2.12b)\qquad \eta = (\theta_2 - \theta_1)/2, \qquad \mu = (\theta_1 + \theta_2)/2.$$

Since both transformations are one to one there will be no loss of generality in taking the observations as Z_0, Z_1, Y_1, Y_2 with joint density

$$(2.13)\quad p(z_0, z_1, y_1, y_2, \mu, \eta) = 2p((z_1 - \mu) - (z_0 - \eta), y_1)p((z_1 - \mu) + (z_0 - \eta), y_2)$$

and writing estimators as $\delta(Z_1 - Z_0, Z_1 + Z_0, Y_1, Y_2)$. (For normal distributions, Z_0 and Z_1 are independent.) The risk function (2.9) becomes, noting that $\varphi(\theta_1, \theta_2)$ is $\mu + |\eta|$,

$$(2.14)\quad R(\mu, \eta, \delta) = \int_{\mathcal{Y}_2} \int_{\mathcal{Y}_1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\delta(z_1 + z_0, z_1 - z_0, y_1, y_2) - (\mu + |\eta|)]^2 \cdot p(z_0, z_1, y_1, y_2, \mu, \eta)\, dz_0\, dz_1 \nu(dy_1)\nu(dy_2).$$

Observe that the estimators (2.10) and (2.11) in this notation are respectively

$$(2.15)\qquad \varphi(Z_0, Z_1) = Z_1 + |Z_0|,$$

and

$$(2.16)\quad \delta^*(Z_1 - Z_0, Z_1 + Z_0, Y_1, Y_2) = Z_1 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\eta| p(Z_0, Z_1, Y_1, Y_2, \mu, \eta)\, d\mu\, d\eta \cdot$$

3. Minimax. In this section, we consider the minimax properties of $\varphi(X_1, X_2)$ and δ^*. We start by proving the minimax property of an estimator for an apparently unrelated problem. Then we reduce our problem, by means of invariance arguments, in order to make use of the above minimax result.

THEOREM 3.1. *Let Z be a real random variable, and let $(\mathcal{Y}, B, \nu)$ be a probability space. Suppose $g(z, y) \geqq 0$, defined on $E_1 \times \mathcal{Y}$ to E_1 is jointly measurable in (z, y) and satisfies (2.7) and (2.8). Let*

$$(3.1)\qquad R = \iint z^2 g(z, y)\, dz \nu(dy) < \infty .$$

Let ω be a location parameter so that for each y, we get the family of densisies $\{g(z - \omega, y), -\infty < \omega < \infty\}$. Then $|Z|$ is a minimax estimator of $|\omega|$, with maximum risk R.

PROOF. The risk for $|Z|$ is

$$\iint (|z| - |\omega|)^2 g(z - \omega, y)\, dz \nu(dy)$$

$$\leqq \iint (|z - \omega|)^2 g(z - \omega, y)\, dz\nu\, d(y) \tag{3.2}$$
$$= \iint (z - \omega)^2 g(z - \omega, y)\, dz\nu(dy) = R.$$

Now suppose $\delta(z, y)$ is an estimator of $|\omega|$, such that for $\epsilon \geqq 0$, and for all ω

$$\iint (\delta(z, y) - |\omega|)^2 g(z - \omega, y)\, dz\nu(dy) \leqq R - \epsilon. \tag{3.3}$$

Then if $\omega \geqq 0$, we have from (3.3)

$$\iint (\delta(z, y) - \omega)^2 g(z - \omega, y)\, dz\nu(dy) \leqq R - \epsilon. \tag{3.4}$$

By a change of variable in (3.4), we see at once that for $\omega \geqq -a$ we have

$$\iint (\delta(z + a, y) - a - \omega)^2 g(z - \omega, y)\, dz\nu(dy) \leqq R - \epsilon. \tag{3.5}$$

Now for the problem of estimating the translation parameter ω with respect to squared error loss it is well known that a minimax invariant estimate with risk R exists. Furthermore, for such a problem, the space of decision functions may be regarded as compact with respect to regular (weak) convergence, after compactification of the action space. (See LeCam (1955), Remark 6.) Hence we may define a sequence of estimators of the form $\delta(z + a) - a$, as $a \to \infty$, such that a subsequence converges weakly to a limit function $\delta^*(z, y)$, which by (3.4) satisfies for all ω,

$$\iint (\delta^*(z, y) - \omega)^2 g(z - \omega, y)\, dz\nu(dy) \leqq R - \epsilon. \tag{3.6}$$

If δ^* does not lie in the original space of decision functions determined before the compactification of the action space, then it could be replaced by a decision procedure in the original decision space whose risk is less than or equal to the risk of δ^* for all ω. This is so since the compactification can be done in accordance with LeCam (1955), Remark 6. Also if the limiting procedure were randomized it could be replaced by a nonrandomized procedure whose risk was at least as good. Thus, since the risk for the minimax procedure for the unrestricted problem is R, it follows from (3.6) that $\epsilon = 0$. Hence from (3.3) there exists no estimator whose risk has a supremum less than R, and the theorem is proved by virtue of (3.2).

REMARK. It appears that the latter portion of the proof of this theorem would follow also from the results of Peisakoff (1950).

To apply Theorem 3.1 to our problem, we now develop some invariance relations satisfied by both $\varphi(X_1, X_2)$ and δ^*. First, these estimators are symmetric in the index i, satisfying $\delta(x_1, x_2, y_1, y_2) = \delta(x_2, x_1, y_2, y_1)$. In the Z notation of Section 2, we shall require then that

$$\delta(Z_1 - Z_0, Z_1 + Z_0, y_1, y_2) = \delta(Z_1 + Z_0, Z_1 - Z_0, y_2, y_1). \tag{3.7}$$

Define,

$$Z_2 = |Z_0| \quad \text{and} \quad \omega = |\eta|. \tag{3.8}$$

If we assume that Y_1 and Y_2 are identical spaces, and use (3.7) and (3.8) in the formula (2.14) for the risk function, we find that $R(\mu, \eta, \delta) = R(\mu, -\eta, \delta) =$

$R(\mu, \omega, \delta)$. Also it is easily seen that we can take the observations as Z_1, Z_2, Y_1, Y_2, write symmetric estimators as $\delta(Z_1, Z_2, Y_1, Y_2) = \delta(Z_1, Z_2, Y_2, Y_1)$ and take the parameter to be estimated as $(\mu + \omega)$. The joint density of Z_1, Z_2 given Y_1, Y_2 is given by

$$\begin{aligned} &p(z_1, z_2, y_1, y_2; \mu, \omega) \\ &\quad = 2[p((z_1 - \mu) - (z_2 - \omega), y_1)p((z_1 - \mu) + (z_2 - \omega), y_2) \\ &\qquad + p((z_1 - \mu) - (z_2 + \omega), y_1)p((z_1 - \mu) + (z_2 + \omega), y_2)], \\ &\qquad\qquad -\infty < z_1 < \infty, \quad 0 \leqq z_2 < \infty. \end{aligned} \tag{3.9}$$

Further, the estimators $\varphi(x_1, x_2)$ and δ^* are translation invariant under the addition of the same constant a to all observations.

We then consider the class of estimators which have the invariance property $\delta(z_1 + a, z_2, y_1, y_2) = \delta(z_1, z_2, y_1, y_2) + a$. (From (2.1), (2.2), (2.12) and (3.8), we see that adding a to each x_{ij} adds a to z_1 and leaves all other y and z values unchanged.) Letting a be $(-z_1)$, we have

$$\delta(z_1, z_2, y_1, y_2) - z_1 = \delta(0, z_2, y_1, y_2) = \gamma(z_2, y_1, y_2), \tag{3.10}$$

so that we shall write a translation-symmetric invariant estimate as $z_1 + \gamma(z_2, y_1, y_2)$ where

$$\gamma(z_2, y_1, y_2) = \gamma(z_2, y_2, y_1). \tag{3.10a}$$

Clearly an estimator is translation-symmetric invariant if and only if it has this form.

Using (3.10), (3.7), (3.8), (3.9), and $\mathcal{Y}_1 \equiv \mathcal{Y}_2$ in (2.14) gives

$$\begin{aligned} R(\mu, \omega, \delta) = R(\omega, \delta) = \iint\int_{-\infty}^{\infty} \int_0^{\infty} [z_1 + \gamma(z_2, y_1, y_2) - \omega]^2 \\ \cdot p(z_1, z_2, y_1, y_2; 0, \omega)\, dz_2\, dz_1 \nu(dy_2)\nu(dy_1). \end{aligned} \tag{3.11}$$

Let

$$\begin{aligned} g_1(z) &= 2\iint\int_0^{\infty} [p(z + (z_2 - \omega), y_1)p(z - (z_2 - \omega), y_2) \\ &\qquad + p(z - (z_2 + \omega), y_1)p(z + (z_2 + \omega), y_2)]\, dz_2 \nu(dy_1)\nu(dy_2) \\ &= 2\iint \{\int_{-\omega}^{\infty} p(z + t, y_1)p(z - t, y_2)\, dt \\ &\qquad + \int_{\omega}^{\infty} p(z - t, y_1)p(z + t, y_2)\, dt\}\, \nu(dy_1)\nu(dy_2) \\ &= 2\iint\int_{-\infty}^{\infty} p(z + t, y_1)p(z - t, y_2)\, dt \nu(dy_1)\nu(dy_2). \end{aligned} \tag{3.12}$$

Write

$$g_2(z, y_1, y_2) = 2 \int_{-\infty}^{\infty} p(t + z, y_1)p(t - z, y_2)\, dt. \tag{3.13}$$

Note that (2.7) and (2.8) imply that

$$\int_{-\infty}^{\infty} z g_2(z, y_1, y_2)\, dz = 0 \quad \text{all} \quad (y_1, y_2). \tag{3.14}$$

Then (3.11) can be expressed as

$$\begin{aligned} R(\omega, \delta) = {} & \int_{-\infty}^{\infty} z^2 g_1(z)\, dz + \iiint_0^{\infty} [\gamma(z, y_1, y_2) - \omega]^2 [g_2(z - \omega, y_1, y_2) \\ & + g_2(z + \omega, y_1, y_2)]\, dz\nu(dy_1)\nu(dy_2) \\ & + 2 \iiint_0^{\infty} [\gamma(z_2, y_1, y_2) - \omega] \{ \int_{-\infty}^{\infty} z_1 p(z_1, z_2, y_1, y_2; \\ & 0, \omega)\, dz_1 \}\, dz_2 \nu(dy_1)\nu(dy_2). \end{aligned} \tag{3.15}$$

When the cross product term in (3.15) vanishes, the risk of an invariant estimator is just a constant plus the risk of an estimator of the restricted parameter ω. That is, the latter risk is the risk for estimating $\omega > 0$, after observing $z > 0$, where the density of z is the mixture $g_2(z - \omega, y_1, y_2) + g_2(z + \omega, y_1, y_2)$. In this important special case, the problems of finding estimators which are minimax, or admissible in the class of invariant estimators reduce to one dimensional problems of estimating ω. In order for the cross product in (3.15) to vanish for all functions $\gamma(\cdot, \cdot, \cdot)$ and values ω, we shall need the inner integral vanishing and this is equivalent to the following condition:

$$\int_{-\infty}^{\infty} zp(z + t, y_1)p(z - t, y_2)\, dz = 0 \text{ for all } -\infty < t < \infty, \text{ all } (y_1, y_2). \tag{3.16}$$

For the expressions (3.13) and (3.16) to be meaningful, the integrals involved must be finite. From the Schwarz inequality we see that this will be the case if both

$$\int_{-\infty}^{\infty} p^2(t, y)\, dt < \infty, \qquad \text{and} \qquad \int_{-\infty}^{\infty} tp^2(t, y)\, dt < \infty. \tag{3.17}$$

Condition (3.17) is concerned with the behavoir of $p(t, y_1)$ near $t = 0$, inasmuch as $\lim_{t\to\infty} tp(t, y) = 0$ is a sufficient condition to assure that the tails of these integrals do not blow up.

In light of the discussion preceding (3.16), we can now state as a corollary of Theorem 3.1,

THEOREM 3.2. *If (3.16) holds, and if (using (3.12) and (3.15))*

$$\begin{aligned} R = {} & \iint_{-\infty}^{\infty} z^2 g_2(z, y)\, dz\nu(dy) + \int_{-\infty}^{\infty} z^2 g_1(z)\, dz \\ & = \iint_{-\infty}^{\infty} x^2 p(x, y)\, dx\nu(dy) < \infty, \end{aligned} \tag{3.18}$$

then $\varphi(X_1, X_2)$ is a minimax estimator of $\varphi(\theta_1, \theta_2)$ with maximum risk of R.

PROOF. In the present notation,

$$\varphi(X_1, X_2) = Z_1 + Z_2 \tag{3.19}$$

so that in view of (3.15), (3.16) and the intervening discussion, to prove the minimax property of (3.19) among symmetric translation invariant estimates, we need only show that if Z has distribution (given Y) of $[g_2(z - \omega, y) + g_2(z + \omega, y)]$ $(z \geqq 0, \omega \geqq 0)$, then Z is a minimax estimator of ω with maximum risk

$$R' = \iint_{-\infty}^{\infty} z^2 g_2(z, y)\, dz\nu(dy) < \infty. \tag{3.20}$$

This last follows from Theorem 3.1 by noting that if X (given Y) has distribution $g_2(x - \theta, y)(-\infty < x, \theta < \infty)$, then $Z = |X|$ has the stated density when

$\omega = |\theta|$. By (3.14) and (3.18), we see that the hypotheses of Theorem 3.1 are satisfied.

The problem is invariant under permutation of the index, and translation of both parameters by a (real). Thus we have demonstrated above that $\varphi(X_1, X_2)$ is minimax among invariant procedures. The conclusion of Theorem 3.2 follows then from the generalization of the well known Hunt-Stein theorem. (See Kiefer (1957).)

We now turn to the estimator $\delta^*(X_1, X_2, Y_1, Y_2)$ given for our purposes by $Z_1 + \gamma^*(Z_2, Y_1, Y_2)$ where

$$(3.21) \quad \gamma^*(Z, Y_1, Y_2) = \int_0^\infty \omega[g_2(Z - \omega, Y_1, Y_2) + g_2(Z + \omega, Y_1, Y_2)]\, d\omega.$$

When (3.16) holds, the risk of δ^* is given by the first two terms of (3.15), say $C + R(\omega, \gamma^*)$. The non-minimax property of δ^* (under (3.16)) will follow from Theorem 3.2 if we show that $R(\omega, \gamma^*) > R'$ (see (3.20)) for some ω. Using

$$(3.22) \qquad \gamma^*(z, y) - z = 2 \int_z^\infty (x - z) g_2(x, y)\, dx > 0, \quad \text{all } z,$$

in (3.15), we find that

$$(3.23) \qquad \begin{aligned} R(0, \gamma^*) = \int \nu(dy)\{&2 \int_0^\infty z^2 g_2(z, y)\, dz \\ &+ 8 \int_0^\infty z g_2(z, y) \int_z^\infty (x - z) g_2(x, y)\, dx\, dz \\ &+ 8 \int_0^\infty g_2(z, y) (\int_z^\infty (x - z) g_2(x, y)\, dx)^2\, dz\} \\ > 2 \int \nu(dy) &\int_0^\infty z^2 g_2(z, y)\, dz. \end{aligned}$$

If $g_2(-z, y) = g_2(z, y)$ (as is the case when $n = 1$ and $f(\cdot)$ is symmetric, or for general n when $p(x, y)$ is symmetric for each y, e.g. when $f(\cdot)$ is normal or uniform), then (3.23) becomes $R(0, \gamma^*) > R'$. We have now demonstrated that $\delta^*(\cdot, \cdot, \cdot, \cdot)$ is not minimax for estimating $\varphi(\theta_1, \theta_2)$ for the $f(\cdot)$'s mentioned above.

We now illustrate that when (3.16) does not hold, $\varphi(X_1, X_2)$ may not be minimax for estimating $\varphi(\theta_1, \theta_2)$. Let $n = 1$ so that $X_i = X_{i1}$ ($i = 1, 2$) and let $f(\cdot)$ be the exponential density given by

$$(3.24) \qquad \begin{aligned} f(x) &= e^{-(x+1)}, & x &\geqq -1, \\ &= 0, & x &< -1. \end{aligned}$$

For the estimator $(\varphi(X_1, X_2) - d)$, a direct calculation gives

$$(3.25) \quad \begin{aligned} R(\theta_1, \theta_2, \varphi(X_1, X_2) - d) &= E(\max(X_1, X_2) - d - \max(\theta_1, \theta_2))^2 \\ &= 1 + (\tfrac{1}{2} - d) e^{-|\theta_2 - \theta_1|} + d^2. \end{aligned}$$

For $d = \frac{1}{2}$, we obtain the constant risk of (5/4), whereas for $d = 0$ (corresponding to $\varphi(X_1, X_2)$), we have risk of $1 + (\frac{1}{2}) \exp(-|\theta_2 - \theta_1|)$ which is $(\frac{3}{2})$ at $|\theta_2 - \theta_1| = 0$ (i.e., $\omega = 0$). Thus $\varphi(X_1, X_2)$ is not minimax. (Incidentally δ^* is not minimax for this example either, having risk exceeding $(\frac{3}{2})$ at $\omega = 0$.)

The above argument also shows that $\varphi(X_1, X_2)$ and δ^* cannot be minimax for general n as well, since the Pitman estimator X_i for the exponential density $f(x - \theta)$ with $f(x)$ given by (3.24) is just (min $(X_{i1}, \cdots, X_{in}) + (n - 1)/n)$. This X_i has an exponential distribution $(ne^{-n(x+1)}, x \geqq -1)$, and is independent of $(Y_1, \cdots, Y_{n-1})$.

REMARKS. 1. First we wish to discuss the relation (3.16) and the conditions under which it holds since that relation is crucial to the results of this section.

Condition (3.16) is satisfied when $n = 1$ for all symmetric densities $f(\cdot)$, and is satisfied for general n if X_i is independent of $Y_i (i = 1, 2)$ and $f(\cdot)$ is symmetric. For example, for arbitrary n, the normal density satisfies (3.16).

It will be noted that (3.16) is equivalent to

$$\text{(3.16a)} \quad E(X_1 + X_2 \mid X_2 - X_1, y) = 0, \quad \text{for all } (X_2 - X_1) \text{ and } y.$$

If X_i is independent of Y_i (as for $n = 1$), Cacoullos (1967) shows that for all distributions, such that all moments are finite, (3.16a) implies that $p(x)$ is symmetric. C. R. Rao and K. Jogdeo, in personal communications, claim that the result can be extended to include all distributions whose characteristic functions are not zero on any interval; but Rao has an example indicating that the result is not always true.

It is also clear that (3.16) holds if $p(x, y)$ is symmetric in x for each fixed y. This will not be true in general when $f(\cdot)$ is symmetric, but does hold for some cases. We shall demonstrate here the symmetry of $p(x, y)$ for each y, for the uniform distribution. Let

$$\text{(3.26)} \qquad f(x) = (\tfrac{1}{2}), \qquad -1 \leqq x \leqq 1.$$

Use the definitions (2.1), (2.2), (2.3), and (2.4) along with

$$\text{(3.27)} \quad Y_0 \equiv 0, \qquad M_n = \max (Y_0, Y_1, \cdots, Y_{n-1}),$$
$$m_n = \min (Y_0, Y_1, \cdots, Y_{n-1}), \qquad R_n(y) = M_n - m_n,$$

to find that

$$\text{(3.28)} \qquad r_0(y) = (\tfrac{1}{2})^n(2 - R_n(y)), \qquad r_1(y) = -(\tfrac{1}{2})(M_n + m_n).$$

From (2.6) we obtain

$$\text{(3.29)} \qquad \begin{aligned} p(x, y) &= [1/(2 - R_n(y))] \qquad \text{for } -1 + (R_n(y)/2) \leqq x \leqq 1 - (R_n(y)/2), \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

which is symmetric in x, so that (3.16) holds.

2. The risk function of $\varphi(X_1, X_2)$ was shown to be the first term in (3.15) $+ R(\omega, Z_2)$, and $R(\omega, Z_2)$ was shown to be bounded by R' (3.20) (for $\omega \geqq 0$). In addition, we indicate the following formula

$$\text{(3.30)} \qquad R(\omega, Z_2) = R' - 4\omega \iint_{\omega}^{\infty} (z - \omega) g_2(z, y)\, dz\nu(dy)$$

from which the property $\lim_{\omega \to \infty} R(\omega, Z_2) = R' = R(0, Z_2)$ is immediate.

3. Additional properties of $\delta^* = Z_1 + \gamma^*$ are: (i) from (3.22) we see that $\delta^*(\cdot, \cdot, \cdot, \cdot) > \varphi(\cdot, \cdot)$ but converges monotonically to $\varphi(\cdot, \cdot)$ as Z_2 increases.

Although we showed $R(0, \gamma^*) > R'$, we can also show that (i) $\lim_{\omega\to\infty} R(\omega, \gamma^*) = R'$, and (ii) $\lim_{N\to\infty} (1/N) \int_0^N R(\omega, \gamma^*)\, d\omega = R'$.

4. Admissibility. In this section we prove the admissibility of the estimator δ^* given in (1.3) and (2.11). The proof follows essentially from results of Stein. In the first lemma below we state Theorem 3.1 of James and Stein (1961) in a form that more closely fits the present problem. This theorem gives a sufficient condition for almost admissibility of an estimator of an arbitrary real valued function $g(\theta_1, \theta_2)$. In Theorem 4.1, we prove under appropriate moment conditions the admissibility of the estimator which equals the *a posteriori* expected value of $g(\theta_1, \theta_2)$ given the observations, and given a uniform generalized prior distribution on (θ_1, θ_2), i.e.

$$\hat{\delta}(X_1, X_2, Y) = \iint g(\theta_1, \theta_2) p(X_1 - \theta_1, X_2 - \theta_2, Y)\, d\theta_1\, d\theta_2 . \tag{4.1}$$

Following this we show that Theorem 4.1 is applicable to δ^* as an estimator of $\varphi(\theta_1, \theta_2)$. We conclude this section by showing that $\varphi(X_1, X_2)$ (see (2.10)) is inadmissible for $\varphi(\theta_1, \theta_2)$ when f is normal.

We now state

LEMMA 4.1. (*Stein*). *If $\delta(X, Y)(X = (X_1, X_2))$ is an estimator of $g(\theta)(\theta = (\theta_1, \theta_2))$ with bounded risk such that for each set C in a denumerable family F of sets whose union is the space of all θ,*

$$\inf_{q\varepsilon S(C)} \left[\iint q(\theta)\, d\theta E_\theta(\delta(X, Y) - \delta_q(X, Y))^2 \Big/ \iint q(\theta)\, d\theta\right] = 0, \tag{4.2}$$

where $S(C)$ is the set of probability densities with respect to Lebesgue measure which are constant (but not 0) on C, and $\delta_q(X, Y)$ is the a posteriori expected value of $q(\theta)$ when $q(\theta)\, d\theta$ is the prior distribution, then $\delta(X, Y)$ is almost admissible with respect to Lebesgue measure.

Before stating Theorem 4.1, let us suppose that C is a disk with center at the origin and radius R, $R \geqq 1$. Then note that (4.2) is implied by

$$\begin{aligned} &\lim_{\sigma\to\infty} (1/\pi_\sigma(0)) \int \nu(dy) \iint dx_1\, dx_2 \\ &\cdot \{[\iint (\delta(x, y) - g(\theta)) p(x_1 - \theta_1, x_2 - \theta_2, y) \pi_\sigma(\theta)\, d\theta]^2 \\ &\qquad /\iint p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta\} = 0, \end{aligned} \tag{4.3}$$

where $\pi_\sigma(\theta)$ is a sequence of densities, with respect to Lebesgue measure, which are constant on the disk C. To see that (4.3) implies (4.2) note that the integral in the numerator of (4.2) is

$$\begin{aligned} &\iint \pi_\sigma(\theta) E_\theta[\delta(x, y) - \iint g(\theta) p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta \\ &\qquad / \iint p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta]^2 \\ &= \iint \pi_\sigma(\theta) \int \nu(dy) \iint p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 \\ &\qquad \cdot \{[\iint (\delta(x, y) - g(\theta)) p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta] \\ &\qquad /[\iint p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta]^2\} \end{aligned} \tag{4.4}$$

$$= \int \nu(dy) \iint dx_1\, dx_2 \cdot \{[\iint (\delta(x, y) - g(\theta))p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta]^2 / \iint p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta\}.$$

Since the sequence $\pi_\sigma(\theta)$ will be constant on the disk C, it is clear from (4.4) that (4.3) implies (4.2). Now we state

THEOREM 4.1. *If the observed variables* (X_1, X_2, Y) *are distributed so that for some* $\theta = (\theta_1, \theta_2)$, $(X_1 - \theta_1, X_2 - \theta_2, Y)$ *has a probability density satisfying* (2.7) *and* (2.8) *and if*

$$\int \nu(dy)[\iint (x_1^2 + x_2^2)|\log^\beta (x_1^2 + x_2^2)|p(x_1, x_2, y)\, dx_1\, dx_2] \cdot [H(y) + R(y)] < \infty, \tag{4.5}$$

where

$$H(y) = \sup_{\theta_1,\theta_2} \iint [g(-x_1, -x_2) - E_\theta\{g(-X_1, -X_2)\}]^2 \cdot p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 \tag{4.6}$$

and

$$R(y) = \sup_\theta R(\theta, y) = \sup_{\theta_1,\theta_2} \iint [\hat{\delta}(x, y) - g(\theta)]^2 \cdot p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 \tag{4.7}$$

then $\hat{\delta}(X, Y)$ *given by* (4.1) *is an admissible estimator of* $g(\theta)$.

PROOF. Since we assume $p(x_1 - \theta_1, x_2 - \theta_2, y)$ is a density for each y, it follows from Stein (1959a), p. 973, that almost admissibility implies admissibility. Hence, by virtue of Lemma 4.1, if we show (4.3), the proof is complete. To demonstrate (4.3) we define π_σ by

$$\begin{aligned} \pi_\sigma(\theta) &= (K_\sigma/\sigma) \log^2 (\sigma^{\frac{1}{2}}/R^{\frac{1}{2}}); & 0 \leqq \theta_1^2 + \theta_2^2 \leqq R^2, \\ &= (K_\sigma/\sigma) \log^2 (\sigma^{\frac{1}{2}}/(\theta_1^2 + \theta_2^2)^{\frac{1}{2}}); & R^2 \leqq \theta_1^2 + \theta_2^2 \leqq M\sigma, \\ &= (K_\sigma/\sigma)[B\sigma/(\theta_1^2 + \theta_2^2)][1/\log^\beta (A((\theta_1^2 + \theta_2^2)/\sigma)^{\frac{1}{2}})]; & \theta_1^2 + \theta_2^2 \geqq M\sigma \end{aligned} \tag{4.8}$$

where $1 < \beta < 1 + \delta$, $0 < M < 1$, $AM^{\frac{1}{2}} > 1$ and these constants and B are chosen so that π is continuous everywhere and continuously differentiable except at $\theta_1^2 + \theta_2^2 = R^2$. Call the inner integral on the left hand side of (4.3) $I(y)$, i.e.

$$I(y) = \iint dx_1\, dx_2\{[\iint (\hat{\delta}(x, y) - g(\theta))p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta]^2 / \iint p(x_1 - \theta_1, x_2 - \theta_2, y)\pi_\sigma(\theta)\, d\theta\}. \tag{4.9}$$

If we can show that

$$I(y) \leqq R(y) \quad \text{for all } y, \tag{4.10}$$

and

$$I(y) \leqq C_1(\log \sigma/\sigma)H(y) \cdot [\iint (x_1^2 + x_2^2)(1 + |\log^\beta (x_1^2 + x_2^2)|p(x_1, x_2, y)\, dx_1\, dx_2], \tag{4.11}$$

whenever

$$\lambda(y) = \iint (x_1^2 + x_2^2) p(x_1, x_2, y)\, dx_1\, dx_2 \leqq C_2\sigma, \tag{4.12}$$

where C_1 and C_2 are positive constants, then we may proceed as in James and Stein (1961), pps. 374 and 375 to complete the proof of the theorem.

Applying the Schwarz inequality to the numerator of (4.9), gives

$$I(y) \leqq \iint \pi_\sigma(\theta) R(\theta, y)\, d\theta \leqq R(y) \tag{4.13}$$

which is (4.10). To establish (4.11), apply (4.1) followed by the Schwarz inequality to the bracketed integral in the numerator of (4.9) to obtain

$$\begin{aligned} &[\iint (\hat{\delta}(x, y) - g(\theta)) p(x_1 - \theta_1, x_2 - \theta_2, y)[\pi_\sigma(\theta) - \pi_\sigma(x)]\, d\theta]^2 \\ &\quad \leqq \iint (\hat{\delta}(x, y) - g(\theta))^2 p(x_1 - \theta_1, x_2 - \theta_2, y)\, d\theta \\ &\qquad \cdot \iint [\pi_\sigma(\theta) - \pi_\sigma(x)]^2 p(x_1 - \theta_1, x_2 - \theta_2, y)\, d\theta. \end{aligned} \tag{4.14}$$

Using (4.1), we find that

$$\begin{aligned} &\iint (\hat{\delta}(x, y) - g(\theta))^2 p(x_1 - \theta_1, x_2 - \theta_2, y)\, d\theta \\ &\quad = \iint g^2(-\theta_1, -\theta_2) p(\theta_1 + x_1, \theta_2 + x_2, y)\, d\theta_1\, d\theta_2 \\ &\qquad - [\iint g(-\theta_1, -\theta_2) p(\theta_1 + x_1, \theta_2 + x_2, y)\, d\theta_1\, d\theta_2]^2. \end{aligned} \tag{4.15}$$

The right side of (4.15) is bounded uniformly in (x_1, x_2) by $H(y)$ (see (4.6)). Using (4.15), (4.14) and (4.6) in (4.9) we now find that

$$\begin{aligned} I(y) &\leqq H(y) \iint dx_1\, dx_2 [\iint [\pi_\sigma(\theta) - \pi_\sigma(x)]^2 p(x_1 - \theta_1, x_2 - \theta_2, y)\, d\theta / \\ &\qquad \iint p(x_1 - \theta_1, x_2 - \theta_2, y) \pi_\sigma(\theta)\, d\theta] \\ &= H(y) \iint dx_1\, dx_2 [\iint [\pi_\sigma(x - \eta) - \pi_\sigma(x)]^2 p(\eta_1, \eta_2, y)\, d\eta / \\ &\qquad \iint p(\eta_1, \eta_2, y) \pi_\sigma(x - \eta)\, d\eta]. \end{aligned} \tag{4.16}$$

Apply the Markov inequality as does Stein (1959b) p. 6 to the integral in the denominator on the right side of (4.16), and change the order of integration to obtain

$$\begin{aligned} I(y) &\leqq C_3 H(y) \iint p(\eta_1, \eta_2, y)\, d\eta_1\, d\eta_2 \\ &\qquad \cdot [\iint [\pi_\sigma(x - \eta) - \pi_\sigma(x)]^2\, dx / \pi_\sigma^*(\|x\| + 2\lambda)], \end{aligned} \tag{4.17}$$

where

$$\lambda = \lambda(y), \quad C_3 > 0, \quad \text{and} \quad \pi_\sigma^*(\|\theta\|) = \pi_\sigma^*((\theta_1^2 + \theta_2^2)^{\frac{1}{2}}) = \pi_\sigma(\theta). \tag{4.18}$$

We now note that the integral in (4.17) has been bounded by Stein (1959b) pps. 7–16 for the case $R = 1$, and that Stein's computations are essentially unaltered by the fact that $R > 1$, so that Stein's bound of

$$(C_4 \log \sigma/\sigma) \iint (x_1^2 + x_2^2)(1 + |\log^\beta (x_1^2 + x_2^2)|) p(x_1, x_2, y)\, dx_1\, dx_2 \tag{4.19}$$

is valid here also. If we use (4.19) in (4.17) we get (4.11) and thus complete the proof of Theorem 4.1.

As a corollary to Theorem 4.1, we can now obtain the admissibility of δ^*. Preliminary to this we obtain a bound on $H(y)$ which we give as

LEMMA 4.2. *Let X_1, X_2 have conditional joint density $p(x_1 - \theta_1, x_2 - \theta_2, y)$ given y, then*

$$(4.20)\quad E[\{\min (X_1, X_2) - E(\min (X_1, X_2))\}^2 \mid \mathcal{Y}] \leqq 2 \iint (x_1^2 + x_2^2) p(x_1, x_2, y)\, dx_1\, dx_2 = 2\lambda(y).$$

PROOF. The proof holds in more generality. For let $\alpha_i = EX_i$. Then

$$\begin{aligned} E[\{\min (X_1, X_2) - E(\min (X_1, X_2))\}^2 \mid \mathcal{Y}] &= E[\{((X_1 + X_2)/2) - (|X_1 - X_2|/2) \\ &\quad - ((\alpha_1 + \alpha_2)/2) + (E|X_1 - X_2|/2)\}^2 \mid \mathcal{Y}] \\ &\leqq 2E[\{(X_1 - \alpha_1 + X_2 - \alpha_2)/2\}^2 \mid \mathcal{Y}] \\ &\quad + 2E[\{(|X_1 - X_2|/2) - E|X_1 - X_2|/2\}^2 \mid \mathcal{Y}] \\ &\leqq 2E[(X_1 - \alpha_1)^2 \mid \mathcal{Y}] + 2E[(X_2 - \alpha_2)^2 \mid \mathcal{Y}] = 2\lambda(y), \end{aligned}$$

where the last inequality follows from the fact that $E(|W| - E|W|)^2 \leqq E(W - EW)^2$.

REMARK. Clearly the bound (4.20) holds also for the variance of $\max (X_1, X_2)$. If the joint density $p(x_1, x_2, y)$ is interchangeable $(p(x_1, x_2, y) = p(x_2, x_1, y))$, then the bound in (4.20) can be sharpened to $\lambda(y)$.

Now we state

THEOREM 4.2. *If the observed variables (X_1, X_2, Y) are distributed so that for some $\theta = (\theta_1, \theta_2)$, $(X_1 - \theta_1, X_2 - \theta_2, Y)$ has a probability density satisfying (2.7) and (2.8) and if*

$$(4.21)\quad \int \lambda(dy)[\iint (x_1^2 + x_2^2)|\log^\beta (x_1^2 + x_2^2)| p(x_1, x_2, y)\, dx_1\, dx_2]^2 < \infty,$$

then $\delta^*(X_1, X_2, Y)$ given by (2.11) is an admissible estimator of $\varphi(\theta_1, \theta_2)$.

PROOF. In view of (4.5), (4.6), (4.7) and the definition of $\lambda(y)$ in (4.12), we need only show that

$$(4.22)\quad H(y) \leqq C_1\lambda(y) \quad \text{and} \quad R(y) \leqq C_2\lambda(y) \quad \text{for some } C_1, C_2 \text{ positive.}$$

The first inequality in (4.22) (with $C_1 = 2$) follows from Lemma 4.2 and the fact that $\max(-a, -b)$ is $(-\min(a, b))$. If we write $\varphi(X_1, X_2) = (X_1 + X_2)/2 + |X_1 - X_2|/2$, use (4.1) in (4.7) we have

$$\begin{aligned} R(\theta, y) &= \iint \{[((x_1 + x_2)/2) - ((\theta_1 + \theta_2)/2)] \\ &\quad + [\iint (|\mu_1 - \mu_2|/2) p(x_1 - \mu_1, x_2 - \mu_2, y)\, d\mu_1\, d\mu_2 \\ (4.23)\qquad &\quad - (|\theta_1 - \theta_2|/2)]\}^2 p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 \\ &\leqq \lambda(y) + 2\iint [\iint \{(|\mu_1 - \mu_2|/2) - (|\theta_1 - \theta_2|/2)\} \\ &\quad \cdot p(x_1 - \mu_1, x_2 - \mu_2, y)\, d\mu_1\, d\mu_2]^2 \\ &\quad \cdot p(x_1 - \theta_1, x_2 - \theta_2, y)\, dx_1\, dx_2 . \end{aligned}$$

If we apply the Schwarz inequality inside the second term on the right side of (4.23) and use once again that $(|a| - |b|)^2 \leqq (a-b)^2 \leqq 2(a^2 + b^2)$ we find that the integral in question is bounded by $4\lambda(y)$. Hence $R(y) \leqq 5\lambda(y)$, which shows (4.22) and completes the proof of Theorem 4.2.

NOTE. In proving Theorem 4.2 we did not use either the independence of the X's or the assumption $\mathcal{Y}_1 \equiv \mathcal{Y}_2$. Thus, the estimator δ^* of (1.3) will be admissible even for unequal sample sizes and for (X_{1i}, X_{2i}) not necessarily independent.

We conclude this section by proving

THEOREM 4.3. *If the density f of the original observations X_{ij} is normal, then $\varphi(X_1, X_2)$ is inadmissible.*

PROOF. We first note that $\varphi(X_1, X_2)$ is a translation-symmetric invariant estimator and from (3.19) $\varphi(X_1, X_2) = Z_1 + Z_2$. Clearly if $Z_1 + Z_2$ is inadmissible among the class of translation-symmetric invariant estimators then it is inadmissible. Since f is normal it follows from the development in Section 3 that the class of admissible translation-symmetric invariant estimates is determined by the class of admissible estimates for estimating ω, when observing (Z_2, Y) and the conditional density of Z_2 given Y is $g_2(z - \omega, y) + g_2(z + \omega, y)$. (See (3.13) for the definition of g_2.) Once again, since f is normal it follows that Z_2 and Y are independent and that g_2 is the normal density with constant and known variance. Without loss of generality we let the variance be 1. Thus, the problem is reduced to showing that the single observation Z on a non-negative random variable with density

$$g(z, \omega) = (2\pi)^{-\frac{1}{2}}(e^{-(z-\omega)^2/2} + e^{-(z+\omega)^2/2}), \qquad z > 0, \quad \omega > 0, \tag{4.24}$$

is an inadmissible estimator of ω when the loss is squared error. Note that the density (4.24) satisfies the conditions of a theorem due to Sacks (1963), Remark 3, p. 766. That is, for $\epsilon > 0$, if we replace $e^{\epsilon\omega}$ by $g(z + \epsilon, \omega)/g(z, \omega)$ in Assumption 3 of Sacks, then this assumption is not violated. Assumption 3 requires that, for each $t > 0$, and each $\epsilon > 0$,

$$\sup_{\omega \geqq 0} (t - \omega)^2 \cdot g(z, \omega)/g(z + \epsilon, \omega) < \infty, \tag{4.25}$$

and

$$\lim_{A \to \infty} \sup_{\omega \geqq A} (t - \omega)^2 \cdot g(z, \omega)/g(z + \epsilon, \omega) = 0. \tag{4.26}$$

If we use (4.24), then we see that conditions (4.25) and (4.26) are satisfied, for we get after simplifying,

$$\begin{aligned} \sup_{\omega \geqq 0} (t - \omega)^2 (e^{\omega z} + e^{-\omega z})/(e^{\omega z} e^{\omega \epsilon} + e^{-\omega z} e^{-\omega \epsilon}) \\ = \sup_{\omega \geqq 0} \{(t - \omega)^2 (1 + e^{-2\omega z})/(e^{\omega \epsilon} + e^{-2\omega z} e^{-\omega \epsilon})\}. \end{aligned} \tag{4.27}$$

Clearly the right hand side of (4.27) is finite, and it is also clear that $\lim_{A \to \infty} \sup_{\omega \geqq A}$ of the bracketed factor on the right hand side of (4.27) is zero. Hence, from Sacks' result, Z will be inadmissible if it is not a generalized Bayes solution. But Z cannot be a generalized Bayes solution, since if it were, there would exist a generalized distribution $\xi(\theta)$ such that

$$z = \int_0^\infty \omega(e^{-(z-\omega)^2/2} + e^{-(z+\omega)^2/2})\, d\xi(\omega) \Big/ \int_0^\infty (e^{-(z-\omega)^2/2} + e^{-(z+\omega)^2/2})\, d\xi(\omega). \tag{4.28}$$

Clearly when z is zero, the right hand side of (4.28) is strictly positive. Hence no $\xi(\omega)$ satisfying (4.28) can exist and the theorem is proved.

REMARK. Note that in this section we proved the admissibility of $Z_1 + \gamma^*(Z, Y_1, Y_2)$ for $\mu + \omega$ when the loss is squared error. The proof required us to view the problem in a two dimensional parameter space. Under appropriate conditions, Z_1 is admissible for μ with respect to squared error loss and $\gamma^*(Z_2, Y_1, Y_2)$ is admissible for ω with respect to squared error loss. The risk of $Z_1 + \gamma^*$ for estimating $\mu + \omega$, under (3.16), is the sum of the risks for estimating the individual parameters. This suggests the following question. Let X_i, $i = 1, 2, \cdots, k$, be independent random variables with distribution $F_i(X, \theta_i)$. Let $\delta_i(x_i)$ be admissible estimators for estimating θ_i with loss function $L(\delta, \theta)$. When will $\delta^* = \sum_{i=1}^k \delta_i(x_i)$ be an admissible estimator for $\theta^* = \sum_{i=1}^k \theta_i$, when the loss function is $L(\delta^*, \theta^*)$? Clearly, this is the case when we estimate normal means with known variances when the loss is squared error. The admissibility result of this section is another example of when this may be done.

Acknowledgment. The authors are grateful to the referee for improvement of the original version of this paper. In particular his efforts are responsible for a stronger statement of Theorem 3.1 and shorter proofs of Theorem 3.1 and Lemma 4.2.

REFERENCES

BECHHOFER, R. E., KIEFER, J. and SOBEL M. (1968). *Sequential Ranking and Identification Procedures*, I.M.S. Monograph Series, Chicago Univ. Press.

CACOULLOS, T. (1967). Characterizations of normality by constant regression of linear statistics on another linear statistic. *Ann. Math. Statist.* **38** 1894–1898.

COHEN, A. (1965). A hybrid problem on the exponential family. *Ann. Math. Statist.* **36** 1185–1206.

FARRELL, R. H. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35** 949–998.

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361–379. Univ. of California Press.

KIEFER, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Statist.* **28** 573–601.

LECAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

PEISAKOFF, M. P. (1950). Transformation parameters. Thesis, Princeton Univ.

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

STEIN, C. (1959a). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30** 970–979.

STEIN, C. (1959b). The admissibility of Pitman's estimator for two location parameters. Technical Report No. 25, Department of Statistics, Stanford Univ.