The Annals of Mathematical Statistics
1970, Vol. 41, No. 6, 2021–2034

ESTIMATION OF THE LAST MEAN OF A MONOTONE SEQUENCE

BY ARTHUR COHEN[1] AND HAROLD B. SACKROWITZ[2]

Rutgers University

1. Introduction and summary. The problem of estimating the larger of two translation parameters, when it is not known which population has the larger parameter, was studied by Blumenthal and Cohen (1968a). Blumenthal and Cohen (1968b) have also investigated various estimators for the larger of two means of two independent normal populations. Problems of estimating the largest of a set of ordered parameters, when it is known which populations correspond to each ordered parameter, have been studied in the discrete case by Sackrowitz (1969). Some questions for the continuous case where one does not know which populations correspond to each ordered parameter have been studied by Dudewicz (1969). In this paper we study various problems of estimating the largest of a set of ordered parameters, when it is known which populations correspond to each ordered parameter. The observed random variables are either normally distributed or are continuous and characterized by a translation parameter. The main portion of the study is devoted to estimating the larger of two normal means when we know which population has the larger mean. Note that in one result below, an example of an estimator which is admissible with respect to a convex loss function, but which is not generalized Bayes is given. We proceed to state the models and list the results.

Let X_i, $i = 1, 2$, be independent normal random variables with means θ_i, and known variances. Without loss of generality we let the variance of X_1 be τ and the variance of X_2 be 1. Assume $\theta_2 \geqq \theta_1$, and consider the problem of estimating θ_2 with respect to a squared error loss function. Let $\delta(X_2)$ be any estimator based on X_2 alone. Consider only those $\delta(X_2)$ which are admissible for estimating θ_2 when X_1 is not observed. The following results are obtained.

(1) If the risk of $\delta(X_2)$ is bounded, then $\delta(X_2)$ is inadmissible. This result can be generalized in a few directions. In fact if θ_i are translation parameters of identical symmetric densities, then for any nonnegative strictly convex loss function $W(\cdot)$, with a minimum at 0, X_2 is an inadmissible estimator. Suitable generalizations for arbitrary sample sizes are given. Another generalization is that if C is any positive constant, then $X_2 \pm C$ is inadmissible as a confidence interval of θ_2.

(2) Let U_τ be the positive solution to the equation $a^2 + (\tau+1)a - \tau = 0$. The quantity U_τ will be such that, $0 \leqq U_\tau < 1$. Then the estimators aX_2, for $0 \leqq a < U_\tau$, are admissible.

It will be shown that no $\delta(X_2)$, such that $\delta(X_2)$ is unbounded below, can be generalized Bayes. Thus this result provides an example of an estimator which is

Received May 14, 1969; revised March 2, 1970.

[1] Research supported by NSF Grant GP-13709.

[2] Research supported by NSF Grant GP-9664.

not generalized Bayes, but which is admissible for the squared error loss function. The results above are also true for estimating the largest of k ordered means with known order, in the case of equal variances. It is interesting that for some $a > 0$, aX_k is admissible, regardless of the size of k. The proof of admissibility of aX_2 uses the methods of Blyth (1951) and Farrel' (1968).

(3) Consider the analogue of the Pitman estimator. That is, the estimator which is generalized Bayes with respect to the uniform prior on the space $\theta_2 \geqq \theta_1$. We prove that this estimator is admissible and minimax.

In the next section the inadmissibility of $\delta(X_2)$, given that the risk of $\delta(X_2)$ is bounded, is proved. In Section 3 we prove aX_2 admissible for $0 \leqq a < U_\tau$. We also extend the result here to the case of estimating the largest of k ordered means with known order, in the case of equal variances. In Section 4 we show that if $\delta(X_2)$ is unbounded below then it cannot be generalized Bayes. Generalizations to the symmetric translation case and arbitrary but equal sample sizes are given in Section 5. The confidence interval result is in Section 6 and the final results on the admissible and minimax property of the analogue to the Pitman estimator are given in Section 7.

Throughout, the letters C, K, and M with or without subscripts are used to denote positive constants, not necessarily the same in all cases. Also the symbols φ and Φ are used to denote the probability density function and cumulative distribution function respectively, of the standard normal.

2. Inadmissibility of $\delta(X_2)$ when its risk is bounded. Let X_i, $i = 1, 2$, be independent normal random variables with means θ_i and known variances. Without loss of generality let the variance of X_1 be τ and the variance of X_2 be 1. Assume $\theta_1 \leqq \theta_2$. The problem is to estimate θ_2 when the loss function is squared error. Let $\delta(X_2)$ be an estimator such that $\delta(X_2)$ is admissible for estimating θ_2 when X_1 is not observed. We will prove that if the risk, $\rho(\delta(X_2); \theta_2)$, is bounded then it is inadmissible. We start by stating a lemma due to L. Brown.

LEMMA 2.1. *Let $\delta(X_2)$ be admissible for θ_2 when X_1 is not observed. Write $\delta(X_2) = X_2 + \alpha(X_2)$. Then $\rho(\delta(X_2); \theta_2)$ is bounded if and only if $\alpha(X_2)$ is bounded.*

PROOF. See Brown (1970), Theorem 3.3.1.

Next we prove

LEMMA 2.2. *Let X_1, X_2 be independent normal variables with means θ_1, θ_2 and variances τ and one respectively. Let $\theta_1 \leqq \theta_2$ and $0 < \tau \leqq 1$. If $\delta(X_2)$ is an inadmissible estimator for θ_2 when $\tau = 1$, then it remains inadmissible when $\tau < 1$.*

PROOF. By hypothesis there exists an estimator $h(X_1, X_2)$ such that

$$\rho(h(X_1, X_2); \theta_1, \theta_2, 1, 1) \leqq \rho(\delta(X_2); \theta_1, \theta_2, 1, 1), \tag{2.1}$$

for every (θ_1, θ_2) with strict inequality for at least one (θ_1, θ_2). (The last two components in $\rho(\cdot\,;\theta_1,\theta_2,\cdot,\cdot)$ represent the variances of X_1, X_2 respectively.) Now let Y be a normal random variable with mean zero, variance $1-\tau$, and

$$\begin{aligned}\rho(h(X_1+Y,X_2);\theta_1,\theta_2,\tau,1) &= \rho(h(X_1,X_2);\theta_1,\theta_2,1,1)\\ &\leqq \rho(\delta(X_2);\theta_1,\theta_2,1,1)\\ &= \rho(\delta(X_2);\theta_1,\theta_2,\tau,1).\end{aligned} \tag{2.2}$$

In (2.2) the next to last step follows from (2.1), while the last step follows since the risk of $\delta(X_2)$ does not depend on the distribution of X_1. Finally if we let

$$h^*(X_1,X_2) = \int_{-\infty}^{\infty} h(X_1+y,X_2)(1/(1-\tau)^{\frac{1}{2}})\varphi(y/(1-\tau)^{\frac{1}{2}})\,dy,$$

it is well known that

$$\rho(h^*(X_1,X_2);\theta_1,\theta_2,\tau,1) \leqq \rho(h(X_1+Y,X_2);\theta_1,\theta_2,\tau,1). \tag{2.3}$$

(See for example Ferguson (1967), Section 2.8.) Thus from (2.2) and (2.3) and the strict inequality in (2.1) we find that h^* is better than $\delta(X_2)$, when the variance of X_1 is τ. This completes the proof of the lemma.

Now we are ready to prove

THEOREM 2.1. *Let $\delta(X_2)$ be admissible for θ_2 when X_1 is not observed. Write $\delta(X_2) = X_2+\alpha(X_2)$. If $\alpha(X_2)$ is bounded above, then $\delta(X_2)$ is inadmissible.*

PROOF. By virtue of Lemma 2.2 it suffices to prove the theorem for all $\tau \geqq 1$. Hence, for the remainder of this theorem, $\tau \geqq 1$. Let

$$\begin{aligned} Z_1 &= [1/(\tau+1)]X_1+[\tau/(\tau+1)]X_2, & Z_2 &= -[1/(\tau+1)]X_1+[1/(\tau+1)]X_2,\\ \eta_1 &= [1/(\tau+1)]\theta_1+[\tau/(\tau+1)]\theta_2, & \eta_2 &= -[1/(\tau+1)]\theta_1+[1/(\tau+1)]\theta_2.\end{aligned} \tag{2.4}$$

It is easy to verify that (Z_1, Z_2) are independent, normal variables with means (η_1, η_2) and variances $[\tau/(\tau+1), 1/(\tau+1)]$. In terms of (Z_1, Z_2) the problem is to estimate $\eta_1+\eta_2$, where $\eta_2 \geqq 0$. Write $\delta(X_2) = X_2+\alpha(X_2) = Z_1+Z_2+\alpha(Z_1+Z_2)$. By hypothesis there is some K such that $\alpha(Z_1+Z_2) \leqq K$. We claim that the estimator $\delta^*(Z_1, Z_2) = Z_1+\max(-K, Z_2)+\alpha(Z_1+Z_2)$, beats $\delta(X_2)$. For note that

$$\begin{aligned}&\rho(\delta;\eta_1,\eta_2)-\rho(\delta^*;\eta_1,\eta_2)\\ &\quad= E[Z_1+Z_2+\alpha-\eta_1-\eta_2]^2-E[Z_1+\max(-K,Z_2)+\alpha-\eta_1-\eta_2]^2\\ &\quad= E[Z_2-\eta_2]^2+2E\alpha(Z_2-\eta_2)-E[\max(-K,Z_2)-\eta_2]^2\\ &\qquad\qquad -2E\alpha[\max(-K,Z_2)-\eta_2]\\ &\quad= E[Z_2-\eta_2]^2-E[\max(-K,Z_2)-\eta_2]^2-2E\alpha[\max(-K,Z_2)-Z_2].\end{aligned} \tag{2.5}$$

Since $[\max(-K, Z_2)-Z_2] \geqq 0$ and $\alpha \leqq K$, it follows from (2.5) that

$$\begin{aligned}&\rho(\delta;\eta_1,\eta_2)-\rho(\delta^*;\eta_1,\eta_2)\\ &\quad\geqq E[Z_2-\eta_2]^2-E[\max(-K,Z_2)-\eta_2]^2-2KE[\max(-K,Z_2)-Z_2]\\ &\quad= E[Z_2-(\eta_2-K)]^2-E[\max(-K,Z_2)-(\eta_2-K)]^2\\ &\quad= E[Z_2+K-\eta_2]^2-E[\max(0,Z_2+K)-\eta_2]^2.\end{aligned} \tag{2.6}$$

Now the right-hand side of (2.6) is greater than 0 for every $\eta_2 \geqq 0$ since it is easily shown that $\max(0, Z_2+K)$ is a better estimator for η_2 than Z_2+K, when $\eta_2 \geqq 0$. This completes the proof of the theorem.

COROLLARY 2.1. *If $\delta(X_2)$ is admissible for θ_2 when X_1 is not observed and $\delta(X_2)$ has bounded risk, then $\delta(X_2)$ is inadmissible.*

PROOF. This corollary is an immediate consequence of Lemma 2.1 and Theorem 2.1.

3. Admissibility of aX_2. In this section we prove that aX_2 is admissible for all a such that, $0 \leqq a < U_\tau$. Here U_τ is the positive solution to $a^2+(\tau+1)a-\tau = 0$. The value U_τ satisfies $0 < U_\tau < 1$. For values of a such that, $U_\tau \leqq a < 1$, the admissibility or inadmissibility of aX_2 is not yet resolved. The number U_τ increases monotonically as τ increases and has limits of one and zero as τ tends to infinity and zero respectively.

After proving aX_2 is admissible ($0 \leqq a < U_\tau$), we finish the section by showing the following: Suppose X_i, $i = 1, 2, \cdots, k$, are independent normal variables with means θ_i, and equal known variances. Then aX_k, for all $0 < a < U_{1/(k-1)}$, is admissible for θ_k, where we assume $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k$.

We now proceed to prove aX_2 admissible. The method of proof is by means of Blyth's ratio, which is defined as follows: Let $\xi(\theta)$ be a two-dimensional proper prior distribution defined on the space $\Omega = \{\theta = (\theta_1, \theta_2) \mid \theta_1 \leqq \theta_2\}$. For any procedure δ, let $r(\xi, \delta)$ be its expected risk. That is, $r(\xi, \delta) = E_\xi \rho(\delta; \theta_1, \theta_2, \tau, 1)$. Denote by δ_ξ, the Bayes procedure when the prior distribution is $\xi(\theta)$. Clearly then the Bayes risk for the prior ξ is $r(\xi, \delta_\xi)$. If we wish to prove δ is admissible, Blyth's method is to use the contrapositive. That is, suppose δ is inadmissible and δ^* is better. Then Blyth's ratio

$$R = [r(\xi,\delta)-r(\xi,\delta^*)]/[r(\xi,\delta)-r(\xi,\delta_\xi)], \tag{3.1}$$

must always satisfy, $0 \leqq R \leqq 1$. If we exhibit some ξ for which R exceeds 1 we have a contradiction.

Now we are ready to prove

THEOREM 3.1. *The estimator aX_2, $0 \leqq a < U_\tau$, is admissible.*

PROOF. The case $a = 0$ is trivial. For $0 < a < U_\tau$, let $\lambda = (1-a)/a$. Also let

$$d\xi_{n,\sigma}(\theta_1, \theta_2) = (K_{n,\sigma}/2\pi)(\sigma^{-\frac{1}{2}} \exp[-(\theta_1+n)^2/2\sigma])(\lambda^{\frac{1}{2}} \exp[-\lambda\theta_2{}^2/2]) \quad \text{for } \theta_1 \leqq \theta_2, \tag{3.2}$$

$$= 0 \quad \text{for } \theta_1 > \theta_2,$$

be a sequence of proper prior densities. Here

$$K_{n,\sigma} = (2\pi)/\int_\Omega (\exp[-(\theta_1+n)^2/2\sigma])(1/\sigma^{\frac{1}{2}})\lambda^{\frac{1}{2}} \exp(-\lambda\theta_2{}^2/2)\, d\theta_1\, d\theta_2,$$

and so $K_{n,\sigma}$ is monotone decreasing in n for each fixed σ, $1 \leqq K_{n,\sigma} \leqq 2$, and $\lim_{n\to\infty} K_{n,\sigma} = 1$. Now suppose δ^* is a better estimator than aX_2. Then by analytic properties of the risk functions of aX_2 and δ^* there exists a rectangle $\infty < a_1 \leqq \theta_1 \leqq 0$, $0 \leqq a_2 \leqq \theta_2 \leqq b_2 < \infty$, in Ω, such that

$$\rho(aX_2; \theta_1, \theta_2, \tau, 1) - \rho(\delta; \theta_1, \theta_2, \tau, 1) \geqq \varepsilon, \tag{3.3}$$

for some $\varepsilon > 0$ and for every θ in the rectangle. Hence for the numerator of Blyth's ratio we have,

$$\begin{aligned} r(\xi_{n,\sigma}, aX_2) - r(\xi_{n,\sigma}, \delta^*) &\geqq (\varepsilon/2\pi) \int_{a_1}^{0} \int_{a_2}^{b_2} \lambda^{\frac{1}{2}} \exp(-\lambda\theta_2{}^2/2) \\ &\qquad \cdot (\sigma^{-\frac{1}{2}} \exp[-(\theta_1+n)^2/2\sigma])\, d\theta_2\, d\theta_1 \\ &= \varepsilon C_1 \int_{\sigma^{-1/2}(a_1+n)}^{n,\sigma^{-1/2}} \varphi(t)\, dt \\ &\geqq (C_2/\sigma^{\frac{1}{2}}) \exp(-n^2/2\sigma), \end{aligned} \tag{3.4}$$

for all $n > -a_1$.

We will demonstrate that the denominator of Blyth's ratio goes to zero faster, as $n \to \infty$, than the last expression on the r.h.s. of (3.4), thus giving us the desired contradiction. If we set $\delta_{n,\sigma} = \delta_{\xi_{n,\sigma}}$, then the denominator is

$$\begin{aligned} r(\xi_{n,\sigma}&, aX_2) - r(\xi_{n,\sigma}, \delta_{n,\sigma}) \\ &= E[aX_2 - \theta_2]^2 - E[\delta_{n,\sigma} - \theta_2]^2 \\ &= E[aX_2 - \delta_{n,\sigma}]^2 + 2E[aX_2 - \delta_{n,\sigma}][\delta_{n,\sigma} - \theta_2] \\ &= E[aX_2 - \delta_{n,\sigma}]^2 + 2E\{E\{(aX_2 - \delta_{n,\sigma})(\delta_{n,\sigma} - \theta_2) \mid X_1, X_2\}\} \\ &= E[aX_2 - \delta_{n,\sigma}]^2, \end{aligned} \tag{3.5}$$

as $\delta_{n,\sigma} = E\{\theta_2 \mid X_1, X_2\}$ by definition. Furthermore $E[aX_2 - \delta_{n,\sigma}]^2 = E[E^2\{\theta_2 - aX_2 \mid X_1, X_2\}]$ so that (3.5) becomes

$$\begin{aligned} r(\xi_{n,\sigma}&, aX_2) - r(\xi_{n,\sigma}, \delta_{n,\sigma}) \\ &= \int\int dx_1\, dx_2 \{[\int_\Omega (\theta_2 - ax_2)(1/\tau^{\frac{1}{2}}) \varphi((x_1-\theta_1)/\tau^{\frac{1}{2}}) \varphi(x_2-\theta_2)\, d\xi_{n,\sigma}(\theta)]^2 / \\ &\qquad \int_\Omega \varphi((x_1-\theta_1)/\tau^{\frac{1}{2}})(1/\tau^{\frac{1}{2}}) \varphi(x_2-\theta_2)\, d\xi_{n,\sigma}(\theta)\}. \end{aligned} \tag{3.6}$$

Hence we must show that the r.h.s. of (3.6) goes to zero faster than the last expression in (3.4).

The r.h.s. of (3.6) can be written as

$$\begin{aligned} &CK_{n,\sigma} \int\int dx_1\, dx_2 \{[(\lambda/(\lambda+1)(\sigma+\tau))^{\frac{1}{2}} \varphi([x_1+n]/(\sigma+\tau)^{\frac{1}{2}}) \varphi(\lambda^{\frac{1}{2}} x_2/(\lambda+1)^{\frac{1}{2}}) \\ &\quad \cdot (\int_{-\infty}^{\infty} \int_{-\infty}^{\theta_2} (\theta_2 - ax_2)(((\lambda+1)(\sigma+\tau)/\sigma\tau)^{\frac{1}{2}} \cdot \varphi([\theta_2 - x_2/(1+\lambda)](\lambda+1)^{\frac{1}{2}}) \\ &\quad \cdot \varphi([\theta_1 + (n\tau - \sigma x_1)/(\sigma+\tau)]((\sigma+\tau)/\sigma\tau)^{\frac{1}{2}})\, d\theta_1\, d\theta_2)]^2 / (\lambda^{\frac{1}{2}}/((\lambda+1)(\sigma+\tau))^{\frac{1}{2}}) \\ &\quad \cdot \varphi([x_1+n]/(\sigma+\tau)^{\frac{1}{2}}) \varphi(\lambda^{\frac{1}{2}} x_2/(\lambda+1)^{\frac{1}{2}}) (\int_{-\infty}^{\infty} \int_{-\infty}^{\theta_2} ((\lambda+1)(\sigma+\tau)/\sigma\tau)^{\frac{1}{2}} \\ &\quad \cdot \varphi([\theta_2 - x_2/(1+\lambda)](\lambda+1)^{\frac{1}{2}}) \varphi([\theta_1 + (n\tau - \sigma x_1)/(\sigma+\tau)]((\sigma+\tau)/\sigma\tau)^{\frac{1}{2}}) \\ &\qquad\qquad \cdot d\theta_1\, d\theta_2)\}. \\ &= CK_{n,\sigma} \int\int dx_1\, dx_2 ((\lambda/(\lambda+1)(\sigma+\tau))^{\frac{1}{2}} \varphi([x_1+n]/(\sigma+\tau)^{\frac{1}{2}}) \varphi(\lambda^{\frac{1}{2}} x_2/(\lambda+1)^{\frac{1}{2}})) \\ &\quad \cdot \{[\int_{-\infty}^{\infty} (\theta_2 - ax_2) \Phi(((\sigma+\tau)/\sigma\tau)^{\frac{1}{2}} [\theta_2 + (n\tau - \sigma x_1)/(\sigma+\tau)])(\lambda+1)^{\frac{1}{2}} \\ &\quad \cdot \varphi((\lambda+1)^{\frac{1}{2}} [\theta_2 - x_2/(1+\lambda)])\, d\theta_2]^2 / \int_{-\infty}^{\infty} \Phi(((\sigma+\tau)/\sigma\tau)^{\frac{1}{2}} \\ &\quad \cdot [\theta_2 + (n\tau - \sigma x_1)/(\sigma+\tau)])(\lambda+1)^{\frac{1}{2}} \varphi((\lambda+1)^{\frac{1}{2}} [\theta_2 - x_2/(1+\lambda)])\, d\theta_2\}. \end{aligned} \tag{3.7}$$

Note the identities

$$(3.8)\qquad (1/d^{\frac{1}{2}})\int_{-\infty}^{\infty}\varphi(z/d^{\frac{1}{2}})\,\varphi(z-v)\,dz = (1/(1+d)^{\frac{1}{2}})\,\varphi(v/(1+d)^{\frac{1}{2}})$$

$$\int_{-\infty}^{\infty}\Phi(z/d^{\frac{1}{2}})\,\varphi(z-v)\,dz = \Phi(v/(1+d)^{\frac{1}{2}}).$$

Now for the term to be squared in the numerator on the r.h.s. of (3.7), let $t=(\lambda+1)^{\frac{1}{2}}[\theta_2-x_2/(1+\lambda)]$ and then start to integrate by parts. To complete the integration in the numerator and to integrate the denominator, let $w=[(t/(1+\lambda)^{\frac{1}{2}})+x_2/(1+\lambda)+(n\tau-\sigma x_1)/(\sigma+\tau)]$, and then use the identities in (3.8) to find that (3.7) becomes

$$(3.9)\qquad CK_{n,\sigma}\int\int dx_1\,dx_2((\lambda/(\lambda+1)(\sigma+\tau))^{\frac{1}{2}}\,\varphi([x_1+n]/(\sigma+\tau)^{\frac{1}{2}})\,\varphi(\lambda^{\frac{1}{2}}x_2/(\lambda+1)^{\frac{1}{2}}))$$

$$\cdot(1/(\lambda+1))^2\{(1+\lambda)(\sigma+\tau)/[(\sigma+\tau)+(1+\lambda)\sigma\tau]\}$$

$$\cdot\varphi^2(((1+\lambda)(\sigma+\tau)/[(\sigma+\tau)+(1+\lambda)\sigma\tau])^{\frac{1}{2}}[x_2/(1+\lambda)+(n\tau-\sigma x_1)/(\sigma+\tau)])/$$

$$\cdot\Phi(((1+\lambda)(\sigma+\tau)/[(\sigma+\tau)+(1+\lambda)\sigma\tau])^{\frac{1}{2}}[x_2/(1+\lambda)+(n\tau-\sigma x_1)/(\sigma+\tau)]).$$

Let

$$(3.10)\qquad k_1^{\frac{1}{2}}=((\lambda+1)(\sigma+\tau)/[(\sigma+\tau)+(1+\lambda)\sigma\tau])^{\frac{1}{2}}$$

and make the change of variables $t=k_1^{\frac{1}{2}}[x_2/(1+\lambda)+(n\tau-\sigma x_1)/(\sigma+\tau)]$, $w=x_2(\lambda/(\lambda+1))^{\frac{1}{2}}$. Then (3.9) becomes

$$CK_{n,\sigma}\int dt[\varphi^2(t)/\Phi(t)](1/(\lambda+1))^2[k_1^{\frac{1}{2}}(\sigma+\tau)^{\frac{1}{2}}/\sigma]\int\varphi(w).$$

$$\varphi([(\sigma+\tau)^{\frac{1}{2}}/\sigma]\{(w/(\lambda(\lambda+1))^{\frac{1}{2}})-(t/k_1^{\frac{1}{2}})+n\})\,dw$$

$$(3.11)\qquad = CK_{n,\sigma}\int dt[\varphi^2(t)/\Phi(t)](1/(\lambda+1))^2(k_1\lambda(\lambda+1))^{\frac{1}{2}}((\sigma+\tau)/\lambda(\lambda+1))^{\frac{1}{2}}(1/\sigma)$$

$$\cdot\int\varphi(w)\,\varphi(((\sigma+\tau)/\lambda(\lambda+1))^{\frac{1}{2}}(1/\sigma)[w-t(\lambda(\lambda+1)/k_1)^{\frac{1}{2}}$$

$$+n(\lambda(\lambda+1))^{\frac{1}{2}}])\,dw$$

$$= CK_{n,\sigma}\int dt[\varphi^2(t)/\Phi(t)](1/(\lambda+1))^2(k_1\lambda(\lambda+1))^{\frac{1}{2}}$$

$$\cdot((\sigma+\tau)/[\sigma^2\lambda(\lambda+1)+(\sigma+\tau)])^{\frac{1}{2}}\,\varphi(((\sigma+\tau)/[\sigma^2\lambda(\lambda+1)+\sigma+\tau])^{\frac{1}{2}}$$

$$\cdot[-t(\lambda(\lambda+1)/k_1)^{\frac{1}{2}}+n(\lambda(\lambda+1))^{\frac{1}{2}}])$$

$$= CK_{n,\sigma}\int dt[\varphi^2(t)/\Phi(t)]((1/(\lambda+1))^2(k_1\lambda(\lambda+1))^{\frac{1}{2}}$$

$$\cdot((\sigma+\tau)/[\sigma^2\lambda(\lambda+1)+(\sigma+\tau)])^{\frac{1}{2}}\,\varphi(k_2^{\frac{1}{2}}[-t+n\,(k_1^{\frac{1}{2}})]),$$

where

$$(3.12)\qquad k_2=\lambda[(\sigma+\tau)+(\lambda+1)\sigma\tau]/[\sigma^2\lambda(\lambda+1)+(\sigma+\tau)].$$

Now let $t^* < 0$, be such that $2t^* = \varphi(t^*)/\Phi(t^*)$. It is easy to verify that such a t^* exists and furthermore for every $t < t^*$, $[\varphi(t)/\Phi(t)] < -2t$. Hence we have that (3.11) is less than or equal to

$$
\begin{aligned}
& CK_{n,\sigma}(1/(\lambda+1))^2\{\textstyle\int_{-\infty}^{t^*} -2k_1(k_2)^{\frac{1}{2}} t\varphi(t)\,\varphi(k_2{}^{\frac{1}{2}}[-t+n(k_1)^{\frac{1}{2}}])\,dt \\
&\qquad + [1/\Phi(t^*)]\textstyle\int_{t^*}^{\infty} k_1(k_2)^{\frac{1}{2}}\,\varphi^2(t)\,\varphi(k_2{}^{\frac{1}{2}}[-t+n(k_1)^{\frac{1}{2}}])\,dt\} \\
& \leqq CK_{n,\sigma}(1/(\lambda+1))^2\{e^{-n^2k_1k_2/2(1+k_2)}\textstyle\int_{-\infty}^{0} -2k_1(k_2)^{\frac{1}{2}}\,t\varphi((1+k_2)^{\frac{1}{2}} \\
&\qquad \cdot[t-n(k_1)^{\frac{1}{2}}k_2/(1+k_2)])\,dt + e^{-n^2k_1k_2/(2+k_2)}[1/\Phi(t^*)]\textstyle\int_{t^*}^{\infty} k_1(k_2)^{\frac{1}{2}} \\
&\qquad \cdot\varphi((2+k_2)^{\frac{1}{2}}[t-n(k_1)^{\frac{1}{2}}k_2/(2+k_2)])\,dt\} \\
& = CK_{n,\sigma}(1/(\lambda+1))^2\{e^{-n^2k_1k_2/2(1+k_2)}[k_1(k_2)^{\frac{1}{2}}(1+k_2)^{\frac{1}{2}}] \\
&\qquad \cdot\textstyle\int_{-\infty}^{-n(k_1)^{1/2}k_2/(1+k_2)^{1/2}} -2[(z/(1+k_2)^{\frac{1}{2}}+n(k_1)^{\frac{1}{2}}k_2/(1+k_2))]\,\varphi(z)\,dz \\
&\qquad + e^{-n^2k_1k_2/(2+k_2)}[1/\Phi(t^*)][k_1(k_2)^{\frac{1}{2}}/(2+k_2)^{\frac{1}{2}}] \\
&\qquad \cdot\textstyle\int_{(t^*-n(k_1)^{1/2}k_2/(2+k_2))(2+k_2)^{1/2}}^{\infty}\varphi(z)\,dz\} \qquad (3.13) \\
& \leqq CK_{n,\sigma}(1/(\lambda+1))^2k_1(k_2)^{\frac{1}{2}}\{1/(1+k_2)\,e^{-n^2k_1k_2/2}+[1/\Phi(t^*)(2+k_2)^{\frac{1}{2}}] \\
&\qquad \cdot e^{-n^2k_1k_2/(2+k_2)}\}.
\end{aligned}
$$

From (3.4) (which is the numerator of Blyth's ratio), (3.7) (which defines the denominator of Blyth's ratio), and the r.h.s. of (3.13), we recognize that the desired contradiction is established if

$$
\begin{aligned}
(3.14)\qquad & (C_2/\sigma^{\frac{1}{2}})\,e^{-n^2/2\sigma}\cdot[(\lambda+1)^2/CK_{n,\sigma}\,k_1(k_2)^{\frac{1}{2}}] \\
&\qquad \cdot\{\exp(n^2k_1k_2/2)/(1+k_2)+\exp(n^2k_1k_2/(2+k_2)/\Phi(t^*)(2+k_2)^{\frac{1}{2}}\},
\end{aligned}
$$

tends to infinity as n tends to infinity. Noting the definitions of k_1 and k_2 in (3.10) and (3.12) respectively, and also noting that $(n^2k_1\,k_2/2) > 2n^2k_1\,k_2/2(2+k_2)$, it follows that (3.14) tends to infinity provided

$$
(3.15)\qquad 2k_1k_2/(2+k_2) > 1/\sigma.
$$

Now from (3.10) and (3.12), (3.15) becomes

$$
(3.16)\quad 2\lambda(\lambda+1)(\sigma+\tau)/[2(\sigma+\tau)+2\sigma^2\lambda(\lambda+1)+\lambda(\sigma+\tau)+\lambda(\lambda+1)\sigma\tau] > 1/\sigma.
$$

If we multiply both sides of (3.16) by σ, and rewrite, we arrive at the desired contradiction provided

$$
(3.17)\ [2\lambda(\lambda+1)\sigma+\lambda(\lambda+1)\tau+\lambda(\lambda+1)\tau] > [2\lambda(\lambda+1)\sigma+\lambda(\lambda+1)\tau+(\lambda+2)(\sigma+\tau)/\sigma].
$$

Clearly (3.17) is true when (3.18) is true, where

$$
(3.18)\qquad \sigma\tau\lambda(\lambda+1) > (\sigma+\tau)(\lambda+2).
$$

From (3.18) it is clear that for any $\tau > 0$, there exists a $\lambda^*(\sigma)$ such that for all $\lambda > \lambda^*(\sigma)$, (3.18) holds. Since $a = 1/(1+\lambda)$, the number U_τ is determined by letting

σ tend to infinity and solving the quadratic equation $\tau\lambda(\lambda+1) = (\lambda+2)$ which in terms of a is $a^2+(\tau+1)a-\tau = 0$. Equivalently U_τ is determined from

$$a^2/(\tau+1)+a-\tau/(\tau+1) = 0. \tag{3.19}$$

From (3.19) it is clear that U_τ tends to 0 and 1 as τ tends to 0 and infinity respectively. Furthermore it is easy to verify that U_τ tends to these limits monotonically. This completes the proof of Theorem 3.1.

Now let X_i, $i = 1, 2, \cdots, k$ be independent normal variables with means θ_i and common known variance taken, without loss of generality, to be 1. Assume $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k$. We wish to estimate θ_k. We show that no matter how large k is, there always exists a set, $0 < a < U$, $0 < U < 1$, such that aX_k is admissible. In order to prove this we will consider a sequence of two-dimensional prior distributions on the subspace where $\theta_1 = \theta_2 = \cdots = \theta_{k-1} = \theta$, and θ_k. The sequence of priors will be proportional to the sequence given in (3.2). Since these priors only put mass on a proper subspace of $\Omega = \{(\theta_1, \theta_2, \cdots, \theta_k): \theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k\}$, we cannot use Blyth's ratio. However we can apply a theorem of Farrell (1968), which is a variation of Blyth's method that overcomes the difficulty mentioned above. We remark that Farrell's theorem would also serve to prove Theorem 3.1. Now we are ready to state

THEOREM 3.2. *The estimator* aX_k, $0 \leqq a < U_{1/(k-1)}$, *is admissible.*

PROOF. Let $\lambda = (1-a)/a$, and consider the sequence of priors $d\xi_n(\theta, \theta_k)$ defined in (3.2), with θ replacing θ_1, and θ_k replacing θ_2. Now define

$$\begin{aligned} d\eta_n(\theta_1, \theta_2, \cdots, \theta_k) &= K\, e^{n^2/2\sigma}\, d\xi(\theta, \theta_k) \quad \text{when} \quad \theta_1 = \theta_2 = \cdots = \theta_{k-1} = \theta \\ &= 0 \qquad\qquad\qquad\qquad\quad \text{otherwise.} \end{aligned} \tag{3.20}$$

Let δ_n be Bayes with respect to η_n and also let E_0 be the set in Ω defined as follows:

$$E_0 = \{\theta_1 = \theta_2 = \cdots = \theta_{k-1} = \theta, \theta_k \mid -1 \leqq \theta \leqq 0, 0 \leqq \theta_2 \leqq 1\}.$$

Now we are ready to apply Farrell's theorem (1968), page 23. To show aX_k admissible we need to show that all the properties in (vii) of Farrell's theorem hold. Obviously a sequence of sets $\{F_n, n \geqq 1\}$ of Ω, $F_n \uparrow \Omega$ exists. The sequences δ_n and η_n are defined above, as is a set E_0, which satisfies condition (viia). Condition (viib) is obvious. To verify condition (viic), note that when $\theta_1 = \theta_2 = \cdots = \theta_{k-1} = \theta$, the joint density of $X_1, X_2, \cdots, X_k$ is

$$\begin{aligned} &(1/(2\pi))^{k/2} \exp -(\tfrac{1}{2}) \textstyle\sum_{i=1}^{k} (X_i-\theta_i)^2 \\ &\quad = (1/(2\pi))^{k/2} \exp -(\tfrac{1}{2})(X_k-\theta_k)^2 \exp -(\tfrac{1}{2}) \textstyle\sum_{i=1}^{k-1} (X_i-\bar{X})^2 \\ &\qquad\qquad \cdot \exp -[(k-1)/2](\bar{X}-\theta)^2, \end{aligned} \tag{3.21}$$

where $\bar{X} = \sum_{i=1}^{k-1} X_i/(k-1)$. Hence δ_n depends only on X_k and $\bar{X}$ which are independent normal variables with means θ_k and θ and variances 1 and $1/(k-1)$ respectively. Clearly $\theta_k \geqq \theta$. Now consider the transformation $Z_1 = \bar{X}$, $Z_i = X_i-\bar{X}$, $i = 2, \cdots, k-1$, $Z_k = X_k$. From this transformation and (3.21) it is clear that to compute the risks for δ_n and aX_k we may ignore Z_i, $i = 2, \cdots, k-1$. In light of this fact and the definition of η_n, we see that the problem reduces to verifying (viic) for a two-dimensional problem in which Z_1 and Z_k are observed. Now we may refer to the proof of Theorem 3.1 with $\tau = 1/(k-1)$, to establish (viic). Finally note that with $\tau = 1/(k-1)$,

$$\begin{aligned}\delta_n &= aX_k + k_1^{\frac{1}{2}}\varphi(k_1^{\frac{1}{2}}[X_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)])/(\lambda+1) \\ &\qquad \cdot \Phi(k_1^{\frac{1}{2}}[(X_k/(1+\lambda))+(n\tau-\sigma\bar{X})/(\sigma+\tau)]) \\ &= aX_k + [k_1^{\frac{1}{2}}/(\lambda+1)]\,v(k_1^{\frac{1}{2}}[X_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)]),\end{aligned} \tag{3.22}$$

where $v(z) = \varphi(z)/\Phi(z)$. Hence the risk of δ_n is

$$\begin{aligned}&E(aX_k-\theta_k)^2 \\ &\quad +2E(aX_k-\theta_k)[k_1^{\frac{1}{2}}/(\lambda+1)]\,v(k_1^{\frac{1}{2}}[X_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)]) \\ &\quad +E[k_1/(\lambda+1)^2]\,v^2(k_1^{\frac{1}{2}}[X_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)]).\end{aligned} \tag{3.23}$$

If we apply the Schwarz inequality to the cross product term in the r.h.s. of (3.23), we see that (viid) will be verified if we can show that for each parameter point in Ω,

$$\lim_{n\to\infty} E\,v^2(k_1^{\frac{1}{2}}[x_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)]) = 0. \tag{3.24}$$

Now $v(z)$ is positive, monotone decreasing, and $v(z) \leqq C_1|z|+C_2$. See Katz (1961). Therefore for every fixed $(X_k, \bar{X})$, v in (3.24) decreases as n increases. Hence

$$\begin{aligned}&v^2(k_1^{\frac{1}{2}}[x_k/(1+\lambda)+(n\tau-\sigma\bar{X})/(\sigma+\tau)]) \\ &\quad \leqq v^2(k_1^{\frac{1}{2}}[x_k/(1+\lambda)+(\tau-\sigma\bar{X})/(\sigma+\tau)]) \\ &\quad \leqq \{C_1[|x_k/(1+\lambda)+(\tau-\sigma\bar{X})/(\sigma+\tau)|]+C_2\}^2.\end{aligned} \tag{3.25}$$

Since the last term in (3.25) is integrable we may apply the dominated convergence theorem to establish (3.24). This verifies (viid) of Farrell's theorem and completes the proof of Theorem 3.2.

We conclude this section with the following remarks. For the two-dimensional problem some unresolved questions are (1) Is aX_2 admissible for all a such that $0 \leqq a < 1$? If this were so when X_1 had an arbitrary variance τ, then it would imply that for the k-dimensional problem, there exists estimators based on X_k alone, which are unbounded below, for which no "substantial" improvement can be made, regardless of how big k is. This follows since the k-dimensional model with equal variances can always be reduced (as in Theorem 3.2) to a two-dimensional model with unequal variances.

(2) Can we characterize all those $\delta(X_2)$ which remain admissible for the two-dimensional problem? In light of Theorem 2.1 and Theorem 3.1, are there any estimators $\delta(X_2) = X_2+\alpha(X_2)$, where $\alpha(X_2)$ is unbounded above, that were admissible for the problem of estimating θ_2 alone, but now are inadmissible?

4. Non-generalized Bayes character of aX_2. In this section the set up is as in Sections 2 and 3. That is X_i, $i = 1, 2$ are independent normal random variables with means θ_i and known variances. Assume $\theta_1 \leqq \theta_2$, and $\delta(X_2)$ is an estimate of θ_2 which is admissible for θ_2 if X_1 is not observed. We show that when X_1 and X_2 are observed, if $\delta(X_2)$ is unbounded below then it cannot be a generalized Bayes estimator of θ_2. We start with the definition of a generalized Bayes procedure as given in Sacks (1963).

DEFINITION. A decision function δ is a generalized Bayes solution (GBS) with respect to a measure ξ on Ω which gives finite measure to bounded subsets of Ω if, for almost all (x_1, x_2) (μ = Lebesgue measure), δ selects (perhaps in a randomized way) a decision among those t's which minimize the (generalized) a posteriori loss

$$\int_\Omega (t-\theta_2)^2 \varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})(1/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta)/ \int_\Omega \varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})(1/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta). \tag{4.1}$$

Furthermore only ξ's for which (4.1) is finite for some t and a.e. (x_1, x_2), are considered.

The requirement that (4.1) be finite for some t, a.e. (x_1, x_2), implies that for this problem, any GBS is the nonrandomized estimator

$$\delta_\xi(X_1, X_2) = \int_\Omega \theta_2\,\varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta)/ \int_\Omega \varphi((X_1-\theta_1/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta),$$

where $\delta_\xi(X_1, X_2) < \infty$, a.e., and also implies that the function

$$g(X_1, X_2) = \int_\Omega \varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d_\xi(\theta) < \infty, \tag{4.2}$$

a.e. It is well known that g is an analytic function and that derivatives of all orders can be computed under the integral signs. (See for example, Ferguson (1967), Section 3.5.) Now we are ready to state

THEOREM 4.1. *If $\delta(X_2)$ is an estimator which is unbounded below then $\delta(X_2)$ cannot be generalized Bayes.*

PROOF. Suppose $\delta(X_2)$ is a GBS for some measure ξ and $\delta(X_2)$ is unbounded below. Then

$$\delta(X_2) = \int_\Omega \theta_2\,\varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta)/g(X_1, X_2), \tag{4.3}$$

and $\lim_{x_2 \to -\infty} \delta(X_2) = -\infty$. Let

$$\delta_1(X_1, X_2) = \int_\Omega \theta_1\,\varphi((X_1-\theta_1)/\tau^{\frac{1}{2}})\,\varphi(X_2-\theta_2)\,d\xi(\theta)/g(X_1, X_2). \tag{4.4}$$

Note that $0 = \partial\delta(X_2)/\partial X_1 = (1/\tau)\,d\delta_1(X_1, X_2)/\partial X_2$. This implies that $\delta_1(X_1, X_2)$ is a function of X_1 only. Furthermore, for any $\xi(\theta)$ defined on Ω, we must get $\delta(X_2) \geqq \delta_1(X_1)$ for every pair (X_1, X_2). This fact implies that $\delta_1(X_1) \equiv -\infty$, since $\delta(X_2)$ is unbounded below. But this contradicts the properties of $g(X_1, X_2)$ and so $\delta(X_2)$ cannot be generalized Bayes.

REMARK. Suppose the problem was to estimate both θ_1 and θ_2 where $\theta_1 \leqq \theta_2$, and the loss was the sum of squared errors. Theorem 4.1 shows that no estimator

(δ_1, δ_2), where δ_i depends on X_i alone, and δ_2 is unbounded below, can be generalized Bayes. Hence by virtue of Sacks' theorem, which applies to the problem of estimating both θ_1 and θ_2, no such (δ_1, δ_2) could be admissible. Yet in Section 3, for estimating θ_2 only, we have seen admissible estimators of θ_2 based on X_2 alone, which are unbounded below.

5. Inadmissibility of X_2 for symmetric translation parameter case. In this section we let X_i, $i = 1, 2$ be independent with identical symmetric densities, each characterized by a translation parameter θ_i. That is, we let the density of X_i be $f(X_i - \theta_i)$ (with respect to Lebesgue measure), with f symmetric. Assume again that $\theta_1 \leqq \theta_2$ and we wish to estimate θ_2. The loss function is any nonnegative strictly convex function $W(t, \theta_2) = W(|t - \theta_2|)$ with a minimum at zero. We prove

THEOREM 5.1. *If f is such that $P(X_1 - X_2 > 0) > 0$ for some (θ_1, θ_2), and if there exists some estimator of θ_2 with finite risk, then X_2 is an inadmissible estimator.*

PROOF. Assume X_2 has finite risk, otherwise the hypotheses provide an immediate proof. We will show that $\delta^*(X_1, X_2) = \max((X_1 + X_2/2), X_2)$ beats X_2. To see this we examine the difference in risks. It is

$$\rho(X_2, \theta_2) - \rho(\delta^*, \theta_2) = \int\int_{\{X_2 < X_1\}} W(X_2, \theta_2) f(X_1 - \theta_1) f(X_2 - \theta_2)\, dX_1\, dX_2$$
$$- \int\int_{\{X_2 < X_1\}} W(\delta^*, \theta_2) f(X_1 - \theta_1) f(X_2 - \theta_2)\, dX_1\, dX_2,$$

as X_2 and δ^* differ only on the set where $X_2 < X_1$. Making the transformation

$$Z_1 = (X_1 + X_2)/2, \qquad Z_2 = (X_2 - X_1)/2 \tag{5.1}$$

and letting

$$\eta_1 = (\theta_1 + \theta_2)/2, \qquad \eta_2 = (\theta_2 - \theta_1)/2 \tag{5.2}$$

we find that

$$\rho(X_2, \theta_2) - \rho(\delta^*, \theta_2)$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{0} [W(Z_1 + Z_2, \eta_1 + \eta_2) - W(Z_1, \eta_1 + \eta_2)] f(Z_2 - Z_1 - \eta_2 + \eta_1) \cdot f(Z_2 + Z_1 - \eta_2 - \eta_1)\, dZ_2\, dZ_1$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{-\eta_2} [W(|t_1 + t_2|) - W(|t_1 - \eta_2|)] f(t_2 - t_1) f(t_2 + t_1)\, dt_2\, dt_1.$$

Now consider the following sets: $S_1 = \{(t_1, t_2) : |t_1| \leqq \eta_2,\ t_2 \leqq -\eta_2\}$, $S_2 = \{(t_1, t_2) : t_1 \leqq -\eta_2, t_2 \leqq -\eta_2\}$, and $S_3 = \{(t_1, t_2) : t_1 \geqq \eta_2, t_2 \leqq -\eta_2\}$. We remark that $\eta_2 \geqq 0$ as $\theta_1 \leqq \theta_2$.

Finally, for $i = 1, 2, 3$, let

$$A_i = \int_{S_i}\int [W(|t_1 + t_2|) - W(|t_1 - \eta_2|)] f(t_2 - t_1) f(t_2 + t_1)\, dt_2\, dt_1.$$

Clearly

$$\rho(X_2, \theta_2) - \rho(\delta^*, \theta_2) = A_1 + A_2 + A_3. \tag{5.3}$$

Since $|t_1+t_2| \geqq |t_1-\eta_2|$ for $(t_1, t_2) \in S_1$, it follows from the properties of $W(\cdot)$ that $A_1 \geqq 0$. If we make a change of variable from t_1 to $-t_1$ in A_3, and use the symmetry property of f (i.e. $f(v) = f(-v)$) we note that

$$A_2+A_3 = \int_{-\infty}^{-\eta_2}\int_{-\infty}^{-\eta_2}[W(|t_1+t_2|)-W(|t_1-\eta_2|)-W(|t_2-t_1|) + W(|t_1+\eta_2|)]f(t_1+t_2)f(t_2-t_1)\,dt_2\,dt_1,$$

which is an integration over S_2.

Let $h(t_1, t_2, \eta_2) = W(|t_1+t_2|)-W(|t_1-\eta_2|)-W(|t_2-t_1|)+W(|t_1+\eta_2|)$. On S_2, if $t_2 > t_1$, then

$$|t_1+\eta_2| = -t_1-\eta_2 \geqq -t_1+t_2 = |t_2-t_1| \quad \text{and}$$

$$|t_1+t_2| = -t_1-t_2 \geqq -t_1+\eta_2 = |t_1-\eta_2|.$$

Hence $W(|t_1+\eta_2|) \geqq W(|t_2-t_1|)$ and $W(|t_1+t_2|) \geqq W(|t_1-\eta_2|)$, which implies $h(t_1, t_2, \eta_2) \geqq 0$.

Say $t_1 > t_2$, then, on S_2, the following are true.

(5.4) (a) $h(t_1, t_2, \eta_2) \geqq W(|t_2-\eta_2|)-W(|t_1-\eta_2|)-W(|t_2-t_1|)+W(0)$

(b) $|t_2-\eta_2|+0 = -t_2+\eta_2 = t_1-t_2-t_1+\eta_2 = |t_2-t_1|+|t_1-\eta_2|.$

For any convex function $W(d)$, if $d_1 \leqq d_2 \leqq d_3 \leqq d_4$ and $d_1+d_4 = d_2+d_3$, then $W(d_4)+W(d_1) \geqq W(d_2)+W(d_3)$. Applying this to (5.4) we find that $h(t_1, t_2, \eta_2) \geqq 0$ on S_2 if $t_1 > t_2$.

Hence $h(t_1, t_2, \eta_2) \geqq 0$ on all of S_2. Thus $A_1+A_2+A_3 \geqq 0$ for all $\theta_1 \leqq \theta_2$ with strict inequality for some $\theta_1 \leqq \theta_2$, by the nature of the integrals, and hence by (5.3) the proof of the theorem is complete.

REMARK 1. Theorem 5.1 would hold with suitable revisions for the case where samples of equal size are taken from each of the two populations. The revisions would be in accordance with the notation and transformations given in Blumenthal and Cohen (1968a) page 504. The symmetry condition on f becomes a condition on $p(x, y)$ defined in that reference and discussed further on page 510.

REMARK 2. All of the results in Sections 2, 3, 4, 6, 7 generalize to the case of arbitrary, not necessarily equal, sample sizes in each population.

6. Inadmissibility of the confidence interval $X_2 \pm C$. In this section we return to the model of Section 2. We say that a confidence interval $I(X_1, X_2)$ for θ_2 is inadmissible if there exists a confidence interval $I^*(X_1, X_2)$, whose length is less than or equal to the length of I for all X_1 and X_2, and whose probability of coverage is greater than or equal to the probability of coverage for the I interval for all (θ_1, θ_2). Strict inequality for coverage probability for at least one (θ_1, θ_2) point, or strict inequality for length for some set of positive measure is also required.

Now we prove

THEOREM 6.1. *The confidence interval $I(X_1, X_2) = X_2 \pm C$ is inadmissible.*

PROOF. Consider the transformation of X's to Z's as given in (2.5). We claim that the interval $I^*(Z_1, Z_2) = I$ if $Z_2 \geqq 0$, (i.e. $I(X_1, X_2) = X_2 \pm C$, if $X_2 \geqq X_1$), and $I^*(Z_1, Z_2) = Z_1 \pm C$ if $Z_2 < 0$, (i.e. $(X_1 + \tau X_2)/(1+\tau) \pm C$ if $X_2 < X_1$), is better than I. Clearly we need only show

$$(6.1) \qquad P\{I^* \supset \theta_2\} - P\{I \supset \theta_2\} \geqq 0,$$

for all (θ_1, θ_2) with strict inequality for at least one (θ_1, θ_2). In terms of Z's and η's the condition (6.1) reduces to

$$(6.2) \quad P\{Z_1 - C \leqq (\eta_1 + \eta_2) \leqq Z_1 + C \mid Z_2 < 0\} \\ - P\{(Z_1 + Z_2) - C \leqq (\eta_1 + \eta_2) \leqq (Z_1 + Z_2) + C \mid Z_2 < 0\} \geqq 0.$$

Noting that the conditional distribution of $(Z_1 + Z_2 - \eta_1 - \eta_2)$ given Z_2 is normal with mean $(Z_2 - \eta_2)$ and variance $(\tau/\tau+1)$, we see that (6.2) is equivalent to

$$(6.3) \quad P\{((\tau+1)/\tau)^{\frac{1}{2}}(-C+\eta_2) \leqq ((\tau+1)/\tau)^{\frac{1}{2}}(Z_1 - \eta_1) \leqq ((\tau+1)/\tau)^{\frac{1}{2}}(C+\eta_2)\} \\ \cdot P\{Z_2 \leqq 0\} - \int_{-\infty}^{0} P\{|U| \leqq C\}\, \varphi((\tau+1)^{\frac{1}{2}}[Z_2 - \eta_2])\, dZ_2 \geqq 0,$$

where U is normal with mean $(Z_2 - \eta_2)$ and variance $(\tau/\tau+1)$. Now the l.h.s. of (6.3) becomes

$$(6.4) \quad \{\Phi(((\tau+1)/\tau)^{\frac{1}{2}}[C+\eta_2]) - \Phi(((\tau+1)/\tau)^{\frac{1}{2}}[-C+\eta_2])\}\, \Phi((\tau+1)^{\frac{1}{2}}\eta_2) \\ - \int_{-\infty}^{0} \{\Phi(((\tau+1)/\tau)^{\frac{1}{2}}[C - Z_2 + \eta_2]) - \Phi(((\tau+1)/\tau)^{\frac{1}{2}}[-C - Z_2 + \eta_2])\} \\ \cdot \varphi((\tau+1)^{\frac{1}{2}}[Z_2 - \eta_2])\, dZ_2 \\ = \{\Phi(((\tau+1)/\tau)^{\frac{1}{2}}[C+\eta_2]) - \Phi(((\tau+1)/\tau^{\frac{1}{2}})[-C+\eta_2])\}\, \Phi(-(\tau+1)^{\frac{1}{2}}\eta_2) \\ - \int_{-\infty}^{-(\tau+1)^{\frac{1}{2}}\eta_2} \{\Phi(((\tau+1)/\tau)^{\frac{1}{2}}[C - t/(\tau+1)^{\frac{1}{2}}]) \\ - \Phi(((\tau+1)/\tau)^{\frac{1}{2}}[-C - t/(\tau+1)^{\frac{1}{2}}])\}\, \varphi(t)\, dt.$$

Using the fact that $\eta_2 \geqq 0$ and using a simple property of the normal distribution we see that the r.h.s. of (6.4) is positive for all η_2. Hence (6.2) > 0, and this proves the theorem.

REMARK. Expression (6.3), and hence the last expression in (6.4), represents the difference in the coverage probabilities between the confidence intervals I and I^*. Note that as $\eta_2 \to \infty$ the r.h.s. of (6.4) tends to zero implying that the infima of the coverage probabilities for both intervals are the same.

7. Minimax and admissible property of analogue of Pitman estimator. In this section we consider the estimator which is generalized Bayes with respect to the uniform prior on the space $\theta_2 \geqq \theta_1$. If we call this estimator δ_p, it is easily shown that

$$(7.1) \qquad \delta_p = X_2 + (\tau^{-\frac{1}{2}} + 1)\, \varphi([x_2 - x_1]/(1+\tau)^{\frac{1}{2}})/\Phi([x_2 - x_1]/(1+\tau)^{\frac{1}{2}}).$$

We prove

THEOREM 7.1. *The estimator δ_p is minimax.*

PROOF. The estimator X_2 is minimax. The proof of this fact would follow essentially the same steps as in Blumenthal and Cohen (1968c), Theorem 3.1, page 519. Hence to prove that δ_p is minimax, it suffices to show that δ_p is better than X_2. Therefore consider again the transformation in (2.5) to Z's and η's. From (7.1) we get $\delta_p = Z_1+Z_2+(1/(\tau+1))\varphi((1+\tau)^{\frac{1}{2}}Z_2)/\Phi((1+\tau)^{\frac{1}{2}}Z_2)$. Since Z_1 and Z_2 are independent we note that δ_p beats $X_2 = Z_1+Z_2$, because $Z_2+(1/(\tau+1))\varphi((1+\tau)^{\frac{1}{2}}Z_2)/\Phi((1+\tau)^{\frac{1}{2}}Z_2)$ is a better estimator for η_2, $\eta_2 \geqq 0$, than is Z_2. This latter fact appears in Katz (1961), page 139. Since δ_p beats X_2 it follows that δ_p is minimax and the theorem is proved.

We conclude this section with

THEOREM 7.2. *The estimator δ_p is admissible.*

PROOF. Consider the problem of estimating both θ_1 and θ_2, where $\theta_1 \leqq \theta_2$, and the loss function is the sum of squared errors. For such a problem let $\delta = (\delta_1, \delta_2)$ be the generalized Bayes estimator for the uniform prior on $\theta_1 \leqq \theta_2$. Clearly $\delta_p = \delta_2$. Blumenthal and Cohen (1968c), page 528 have proven that δ is admissible for (θ_1, θ_2), $\theta_1 \leqq \theta_2$, when the variance of X_1 equals the variance of X_2. However it can be shown that their proof suffices to prove δ admissible when X_1 has variance τ and X_2 has variance 1. Now if δ^* beats δ_p for θ_2, then (δ_1, δ^*) would beat $\delta = (\delta_1, \delta_p)$ for (θ_1, θ_2). But this would contradict the admissibility of δ. Hence we conclude that δ_p is admissible and the proof of the theorem is complete.

REFERENCES

[1] BLUMENTHAL, S. and COHEN, A. (1968a). Estimation of the larger translation parameter. *Ann. Math. Statist.* **39** 502–516.

[2] BLUMENTHAL, S. and COHEN, A. (1968b). Estimation of the larger of two normal means. *J. Amer. Statist. Assoc.* **63** 861–876.

[3] BLUMENTHAL, S. and COHEN, A. (1968c). Estimation of two ordered translation parameters. *Ann. Math. Statist.* **39** 517–530.

[4] BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Statist.* **22** 22–42.

[5] BROWN, L. (1970). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. Submitted to *Ann. Math. Statist.*

[6] DUDEWICZ, E. J. (1969). Estimation of ordered parameters. Technical Report, Department of Operations Research, Cornell University.

[7] FARRELL, R. H. (1968). On a necessary and sufficient condition for admissibility of estimators when strictly convex loss is used. *Ann. Math. Statist.* **39** 23–28.

[8] FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.

[9] KATZ, M. W. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.* **32** 136–142.

[10] SACKROWITZ, H. B. (1969). Estimation for monotone parameter sequences: the discrete case. *Ann. Math. Statist.* **41** 609–620.

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