The Annals of Statistics
1987, Vol. 15, No. 3, 1245–1256

ADMISSIBLE MINIMAX ESTIMATION OF A COMMON MEAN OF TWO NORMAL POPULATIONS

BY TATSUYA KUBOKAWA

University of Tsukuba

Consider the problem of estimating the common mean μ of two normal populations with unknown variances σ_1^2 and σ_2^2 under the quadratic loss $(\hat{\mu} - \mu)^2/\sigma_1^2$. A family of minimax estimators with smaller risk than the sample mean in the first population is given, out of which admissible minimax estimators are developed. A class of better estimators of μ under squared-error loss, which is wider than found by Bhattacharya, is obtained.

1. Introduction. Let $(X_1, \dots, X_m)$ and $(Y_1, \dots, Y_n)$ be independent random samples from two normal populations with common unknown mean μ and unknown variances σ_1^2 and σ_2^2, respectively. Also, let $m \geq 2$ and $n \geq 2$. We want to estimate μ by an estimator $\hat{\mu}$ under the quadratic loss

$$L(\hat{\mu}; \mu, \sigma_1^2) = (\hat{\mu} - \mu)^2/\sigma_1^2. \tag{1.1}$$

An estimator will be evaluated by its risk function $E_{\mu, \sigma_1^2, \sigma_2^2}[L(\hat{\mu}; \mu, \sigma_1^2)]$. The justification of the problem discussed here is given in the introduction of Brown and Cohen (1974).

Concerning minimax estimation of the common mean under the loss (1.1), Zacks (1970) showed that the sample mean $\overline{X}$ in the first population is minimax with a constant risk and Cohen and Sackrowitz (1974) obtained minimax estimators better than $\overline{X}$. Some classes of the combined estimators better than $\overline{X}$ have been given by Graybill and Deal (1959), Brown and Cohen (1974) and Khatri and Shah (1974) and have been extended by Bhattacharya (1978, 1980) and Kubokawa (1987). These better estimators are, of course, also minimax for the loss (1.1). On the other hand, the question of admissibility of the well-known estimators, including the Graybill–Deal estimator, is still open as stated in Sinha and Mouqadem (1982) who have discussed it in a restricted class of estimators. Of great interest is finding an estimator which is both minimax and admissible: A much broader class of minimax estimators is desirable.

The object of the present paper is to develop an admissible minimax estimator of μ with respect to the loss (1.1). In Section 2, we provide a family of minimax estimators better than $\overline{X}$. This family includes not only combined estimators in the extended classes of Bhattacharya (1980) and Kubokawa (1987) but also new types of combined estimators. In Section 3, we shall look for an admissible estimator in this family. In particular, we consider the Bayes equivariant estimator which minimizes the Bayes risk among the location equivariant estimators. This Bayes equivariant estimator belongs to our family of minimax estimators and can be proved to be admissible based on Brown

Received May 1986; revised January 1987.
AMS 1980 *subject classifications*. Primary 62F10; secondary 62C15.
Key words and phrases. Common mean, admissible minimax estimator, unbiased estimator.

(1966). This shows that it is admissible and minimax, although it appears to be computationally somewhat complicated. It is also remarked that sample sizes m and n need to be large in order for the estimator to have these desirable properties. These facts show that our estimator may be practically of limited value. As noted in Section 2, estimators included in this family are better than $\bar{X}$ if squared-error loss or squared error divided by a positive function of (σ_1^2, σ_2^2) is taken.

2. A family of minimax estimators. Let $\bar{X} = \sum_{i=1}^m X_i/m$, $S_1 = \sum_{i=1}^m (X_i - \bar{X})^2/m$ and let $\bar{Y}, S_2$ be defined similarly. For nonnegative constants a, b and c, consider estimators of the form

$$\hat{\mu}_\phi(a, b, c) = \bar{X} + \frac{a}{1 + R\phi\left(S_1, S_2, (\bar{X} - \bar{Y})^2\right)}(\bar{Y} - \bar{X}), \tag{2.1}$$

where $R = \{bS_2 + c(\bar{X} - \bar{Y})^2\}/S_1$ and ϕ is a positive function. By the symmetry of the distribution of $\bar{Y} - \bar{X}$, it is seen that $\hat{\mu}_\phi(a, b, c)$ is an unbiased estimator of μ.

THEOREM 2.1. *Assume that* $0 < a \le 2$ *and* $b \ge c \ge 0$, $b > 0$. *Then the estimator* $\hat{\mu}_\phi(a, b, c)$ *given by* (2.1) *is minimax relative to the loss* (1.1) *if the following conditions hold for some* $t \ge 0$:

(a) $n > 2t + 5$ *if* $c = 0$, *or* $n > 2t + 2$ *if* $c > 0$.

(b) $\phi(S_1, S_2, (\bar{X} - \bar{Y})^2)/R^t$ *is nondecreasing in* S_1 *and nonincreasing in* S_2.

(c) *When* $c > 0$, $\phi(S_1, S_2, (\bar{X} - \bar{Y})^2)/R^t$ *is nonincreasing in* $(\bar{X} - \bar{Y})^2$. *In addition, for any* $d_2 > d_1 > 0$, *there exists a function* $\Phi(S_1, S_2; d_1, d_2)$ *independent of* $(\bar{X} - \bar{Y})^2$ *such that, given* $(\bar{X} - \bar{Y})^2 = u$,

$$\frac{d}{du}\left\{\frac{(bS_2 + cu)^t}{\phi(S_1, S_2, u)}\right\} \le \Phi(S_1, S_2; d_1, d_2),$$

for all u *in the interval* (d_1, d_2), *and* $E[\Phi(S_1, S_2; d_1, d_2)|S_1] < \infty$, *where* $E[\cdot|S_1]$ *designates the conditional expectation of the term in brackets, given* S_1.

(d) $$\frac{1}{\phi\left(S_1, S_2, (\bar{X} - \bar{Y})^2\right)} \le \frac{2(n - 2t - 2)}{a(m + 2t + 1)}u(b, c; t),$$

where

$$u(b, c; t) = \frac{E\left[\{b(1 - W) + cW\}^{-t-1}\right]}{E\left[\{b(1 - W) + cW\}^{-t-2}\right]} \tag{2.2}$$

and W *has a beta distribution with parameters* $(3/2, (n - 1)/2)$.

When squared-error loss $(\hat{\mu} - \mu)^2$ or squared error divided by a positive function of (σ_1^2, σ_2^2) is adopted instead of (1.1), Theorem 2.1 remains true just by replacing "minimax" with "better than $\bar{X}$." Hence, our class of better estimators is wider than Bhattacharya (1980) and Kubokawa (1987). Several examples are given at the end of this section. In Section 3, Theorem 2.1 is applied for the particular value $t = 0$. However, treating the case of $t \geq 0$ gives the wider family of minimax estimators. To prove Theorem 2.1, we need the following lemma due to Bhattacharya (1984), Theorem 2.1.

LEMMA 2.1. *Let u, v and w be functions of random variables $X_1, \dots, X_k$ such that v is positive with a finite expectation and $E[vw] > 0$. Let $f_i(X_i|X_1, \dots, X_{i-1}) = E[vw|X_1, \dots, X_i]/E[v|X_1, \dots, X_i]$ and $g_i(X_i|X_1, \dots, X_{i-1}) = E[u|X_1, \dots, X_i]/E[v|X_1, \dots, X_i]$. Then*

$$E[uw]/E[vw] \geq E[u]/E[v],$$

if, for all $1 \leq i \leq k$, either both $f_i(X_i|X_1, \dots, X_{i-1})$ and $g_i(X_i|X_1, \dots, X_{i-1})$ are monotonic in the same direction with respect to X_i, or f_i or g_i is a constant with respect to X_i. This inequality is reversed if, for all $1 \leq i \leq k$, either $f_i(X_i|X_1, \dots, X_{i-1})$ and $g_i(X_i|X_1, \dots, X_{i-1})$ are monotonic in opposite directions with respect to X_i, or f_i or g_i is a constant with respect to X_i.

Although Bhattacharya (1984) does not state that $f_i(X_i|X_1, \dots, X_{i-1})$ or $g_i(X_i|X_1, \dots, X_{i-1})$ may be a constant in order for Lemma 2.1 to hold, his proof covers this case. When f_i (resp. g_i) is a constant with respect to X_i, g_i (resp. f_i) may be an arbitrary function.

PROOF OF THEOREM 2.1. Let $\rho = m\sigma_2^2/(n\sigma_1^2)$ and let T be a random variable such that $T(\sigma_1^2/m + \sigma_2^2/n)^{-1}$ is distributed as a chi-square variate with 3 degrees of freedom and independent of (S_1, S_2). Denote $\bar{\phi} = \phi(S_1, S_2, T)$ and $\bar{R} = (bS_2 + cT)/S_1$. Both Brown and Cohen (1974) and Khatri and Shah (1974) have shown that the estimator $\hat{\mu}_\phi(a, b, c)$ has a smaller risk than $\bar{X}$ if and only if

$$\inf_{\rho > 0} \left\{ \frac{1}{1+\rho} \frac{E\left[(1 + \bar{R}\bar{\phi})^{-1}\right]}{E\left[(1 + \bar{R}\bar{\phi})^{-2}\right]} \right\} \geq \frac{a}{2}. \tag{2.3}$$

As proved in Zacks (1970), $\bar{X}$ is a minimax estimator with a constant risk under the loss (1.1). As a result, (2.3) becomes a necessary and sufficient condition for the minimaxity of $\hat{\mu}_\phi(a, b, c)$. Now, letting $\Theta = \rho/(1 + \rho)$ and using Theorem 2.2 in Bhattacharya (1984) gives

$$\begin{aligned} \frac{1}{1+\rho} \frac{E\left[(1 + \bar{R}\bar{\phi})^{-1}\right]}{E\left[(1 + \bar{R}\bar{\phi})^{-2}\right]} &= \frac{E\left[\{(1 - \Theta) + \Theta\bar{R}\bar{\phi}/\rho\}^{-1}\right]}{E\left[\{(1 - \Theta) + \Theta\bar{R}\bar{\phi}/\rho\}^{-2}\right]} \\ &\geq \min\left\{1, \frac{1}{\rho} \frac{E\left[(\bar{R}\bar{\phi})^{-1}\right]}{E\left[(\bar{R}\bar{\phi})^{-2}\right]}\right\}. \end{aligned} \tag{2.4}$$

[Under condition (a), all expectations in (2.4) are finite.] From condition (d), the r.h.s. of (2.4) is bounded below by

$$\min\left\{1, \frac{a(m+2t+1)}{2(n-2t-2)u(b,c;t)\rho}\frac{E[\bar{R}^{-1}\bar{\phi}^{-1}]}{E[\bar{R}^{-2}\bar{\phi}^{-1}]}\right\}. \tag{2.5}$$

Here, if the inequalities

$$\frac{E[\bar{R}^{-1}\bar{\phi}^{-1}]}{E[\bar{R}^{-2}\bar{\phi}^{-1}]} \geq \frac{E[\bar{R}^{-t-1}]}{E[\bar{R}^{-t-2}]}, \qquad \text{for any } \rho > 0, \tag{2.6}$$

$$\frac{1}{\rho}\frac{E[\bar{R}^{-t-1}]}{E[\bar{R}^{-t-2}]} \geq \frac{n-2t-2}{m+2t+1}u(b,c;t), \quad \text{for any } \rho > 0, \tag{2.7}$$

are valid, then combining (2.4), (2.5), (2.6) and (2.7) yields

$$\inf_{\rho>0}\left\{\frac{1}{1+\rho}\frac{E\left[(1+\bar{R}\bar{\phi})^{-1}\right]}{E\left[(1+\bar{R}\bar{\phi})^{-2}\right]}\right\} \geq \min\left(1, \frac{a}{2}\right), \tag{2.8}$$

which implies (2.3) since $0 < a \leq 2$. Hence, in order to complete the proof, we only need to show (2.6) and (2.7).

To prove (2.6), we use Lemma 2.1. Let

$$\begin{aligned}
f_1(S_1) &= E[\bar{R}^{-2}\bar{\phi}^{-1}|S_1]/E[\bar{R}^{-t-2}|S_1],\\
g_1(S_1) &= E[\bar{R}^{-t-1}|S_1]/E[\bar{R}^{-t-2}|S_1],\\
f_2(T|S_1) &= E[\bar{R}^{-2}\bar{\phi}^{-1}|S_1, T]/E[\bar{R}^{-t-2}|S_1, T],\\
g_2(T|S_1) &= E[\bar{R}^{-t-1}|S_1, T]/E[\bar{R}^{-t-2}|S_1, T],
\end{aligned}$$

and $f_3(S_2|S_1, T)$ and $g_3(S_2|S_1, T)$ be defined similarly. Obviously, both $f_1(S_1)$ and $g_1(S_1)$ are nonincreasing in S_1 by condition (b). Next, for $c = 0$, $g_2(T|S_1)$ is a constant with respect to T. For $c > 0$, both $f_2(T|S_1)$ and $g_2(T|S_1)$ are nondecreasing in T. In fact, $f_2(T|S_1) = E[(bS_2 + cT)^{-t-2}\{(bS_2 + cT)^t\bar{\phi}^{-1}\}|S_1, T]/\{S_1^t E[(bS_2 + cT)^{-t-2}|S_1, T]\}$ and its derivative with respect to T can be written

$$\begin{aligned}
&C_1(S_1,T)E\left[\frac{1}{(bS_2+cT)^{t+2}}\frac{d}{dT}\left\{\frac{(bS_2+cT)^t}{\bar{\phi}(S_1,S_2,T)}\right\}\middle| S_1, T\right]\\
&\quad + C_2(S_1,T)\left\{\frac{E\left[(bS_2+cT)^{-2}\bar{\phi}^{-1}|S_1,T\right]}{E\left[(bS_2+cT)^{-3}\bar{\phi}^{-1}|S_1,T\right]}\right.\\
&\qquad\left. - \frac{E\left[(bS_2+cT)^{-t-2}|S_1,T\right]}{E\left[(bS_2+cT)^{-t-3}|S_1,T\right]}\right\},
\end{aligned} \tag{2.9}$$

where $C_1(S_1, T)$ and $C_2(S_1, T)$ are positive functions. [From conditions (c) and (d), interchange of integration and differentiation is permissible.] By condition (c), the first term in (2.9) is positive. In the second term, both $(bS_2 + cT)^t \bar{\phi}^{-1}$ and $bS_2 + cT$ are nondecreasing in S_2. Thus, from Lemma 2.1, (2.9) is nonnegative, so that $f_2(T|S_1)$ is nondecreasing in T. On the other hand, a similar argument gives that $g_2(T|S_1)$ is nondecreasing in T. From condition (b), it is clear that both $f_3(S_2|S_1, T)$ and $g_3(S_2|S_1, T)$ are nondecreasing in S_2. In this way, we apply Lemma 2.1 to get (2.6).

To prove (2.7), we first express the random variable $\overline{R}$ by other random variables whose distributions are independent of unknown parameters. Let

$$F = \frac{nS_2/\sigma_2^2 + T\left(\sigma_1^2/m + \sigma_2^2/n\right)^{-1}}{mS_1/\sigma_1^2},$$

$$W = \frac{T\left(\sigma_1^2/m + \sigma_2^2/n\right)^{-1}}{nS_2/\sigma_2^2 + T\left(\sigma_1^2/m + \sigma_2^2/n\right)^{-1}}.$$

It is seen that $\{(m-1)/(n+2)\}F$ has an F distribution with $(n+2, m-1)$ degrees of freedom and independent of W, which has a beta distribution with parameters $(3/2, (n-1)/2)$. Note that $F = S_2/(\rho S_1) + FW$. Then $\overline{R}/\rho$ becomes

$$\begin{aligned}\overline{R}/\rho &= \{b\rho F(1-W) + c(1+\rho)FW\}/\rho \\ &= \{b(1-W) + cW + cW/\rho\}F.\end{aligned}$$

By using this expression, the l.h.s. of (2.7) is rewritten as

$$\frac{1}{\rho}\frac{E[\overline{R}^{-t-1}]}{E[\overline{R}^{-t-2}]} = \frac{E\left[\{b(1-W) + cW + cW/\rho\}^{-t-1}\right]E[F^{-t-1}]}{E\left[\{b(1-W) + cW + cW/\rho\}^{-t-2}\right]E[F^{-t-2}]},$$

since F is independent of W. It is observed that $E[F^{-t-1}]/E[F^{-t-2}] = (n-2t-2)/(m+2t+1)$, so that (2.7) is equivalent to $E[\{b(1-W) + cW + cW/\rho\}^{-t-1}]/E[\{b(1-W) + cW + cW/\rho\}^{-t-2}] \geq u(b, c; t)$. Since $b(1-W) + cW \leq b(1-W) + cW + cW/\rho$, it suffices to show that

$$\begin{aligned}&\frac{E\left[\{b(1-W) + cW + cW/\rho\}^{-t-1}\right]}{E\left[\{b(1-W) + cW + cW/\rho\}^{-t-1}\{b(1-W) + cW\}^{-1}\right]} \\ &\qquad \geq \frac{E\left[\{b(1-W) + cW\}^{-t-1}\right]}{E\left[\{b(1-W) + cW\}^{-t-2}\right]}.\end{aligned} \tag{2.10}$$

Evidently, both $\{b(1-W) + cW\}^{t+1}/\{b(1-W) + cW + cW/\rho\}^{t+1}$ and $b(1-W) + cW$ are nonincreasing in W for $b \geq c$. Hence, (2.10) follows from Lemma 2.1, which establishes (2.7). Theorem 2.1 is completely proved. □

For some particular b and c we can obtain exact expressions and a useful inequality for $u(b, c; t)$ defined by (2.2).

LEMMA 2.2. (i) *For* $b = c$, $u(b, b; t) = b$. *For* $c = 0$, $u(b, 0; t) = b(n - 2t - 5)/(n - 2t - 2)$.

(ii) *For* $b > c > 0$, $u(b, c; t)$ *is exactly expressed by*

$$(2.11) \qquad u(b, c; t) = \frac{{}_2F_1(t + 1, 3/2; (n+2)/2; (b-c)/b)}{{}_2F_1(t + 2, 3/2; (n+2)/2; (b-c)/b)} b,$$

where ${}_2F_1$ *is the hypergeometric function. Further,*

$$(2.12) \qquad u(b, c; t) \geq c\left\{1 + \frac{(n-1)(b-c)}{(n+2)b}\right\}.$$

PROOF. First, (i) easily follows since

$$E\left[(1 - W)^{-r}\right] = B(3/2, (n-1)/2 - r)/B(3/2, (n-1)/2) \quad \text{for } r < (n-1)/2,$$

where $B(\, , \,)$ denotes the beta function. Second, (2.11) can be derived by noting that $E[(1 - xW)^{-t-1}] = \sum_{k=0}^{\infty}\{(t+1)_k/k!\}E[W^k]x^k$ for $0 < x < 1$ and $(t+1)_k = (t+1)(t+2) \cdots (t+k)$, and $E[W^k] = B(3/2 + k, (n-1)/2)/B(3/2, (n-1)/2)$. From (2.11), we can get (2.12) by using the following relations in turn: For any real values α, β, γ and $0 < x < 1$,

$$(2.13) \qquad {}_2F_1(\alpha, \beta; \gamma; x) = (1 - x)^{\gamma - \alpha - \beta} {}_2F_1(\gamma - \alpha, \gamma - \beta; \gamma; x)$$

[this is applied to both the numerator and the denominator in (2.11)],

$$(2.14) \quad {}_2F_1(\alpha + 1, \beta; \gamma; x) = {}_2F_1(\alpha, \beta; \gamma; x) + \frac{\beta}{\gamma} x \, {}_2F_1(\alpha + 1, \beta + 1; \gamma + 1; x),$$

and for positive α, β, γ satisfying $(\alpha + \beta + 1)\gamma \geq \alpha\beta$,

$$(2.15) \qquad {}_2F_1(\alpha + 1, \beta + 1; \gamma + 1; x) \geq {}_2F_1(\alpha, \beta; \gamma; x).$$

Equation (2.13) is from Exton (1978) and (2.14) is obtained just by rearrangement of the coefficients in the infinite series on the l.h.s. Inequality (2.15) is also shown by evaluation of each term in the infinite series. Thus, Lemma 2.2 is established. □

The weaker version given by the r.h.s. of (2.12) is noted to make condition (d) of Theorem 2.1 simple and useful. Some examples of minimax estimators based on Theorem 2.1 are given next.

EXAMPLE 2.1. Setting $\phi = 1 + d/\{bS_2 + c(\overline{X} - \overline{Y})^2\}$ in (2.1) for $d \geq 0$, we have

$$\hat{\mu}_1(a, b, c, d) = \overline{X} + \frac{aS_1}{S_1 + bS_2 + c(\overline{X} - \overline{Y})^2 + d}(\overline{Y} - \overline{X}),$$

which includes the estimator $\hat{\mu}_1(1, (m-1)/(n-1), 0, 0)$ of Graybill and Deal (1959); $\hat{\mu}_1(a, (m-1)/(n+2), (m-1)/(n+2), 0)$ of Brown and Cohen (1974);

$\hat{\mu}_1(1, b, 0, 0)$ and $\hat{\mu}_1(1, b, b, 0)$ of Khatri and Shah (1974); $\hat{\mu}_1(a, b, 0, 0)$ and $\hat{\mu}_1(a, b, b, 0)$ of Bhattacharya (1980) and $\hat{\mu}_1(a, b, c, 0)$ with $b \geq c \geq 0$ of Kubokawa (1987). Then Theorem 2.1 and Lemma 2.2 with $t = 0$ imply that $\hat{\mu}_1(a, b, c, d)$ is minimax, either if $a \leq 2\min\{1, b(n-5)/(m+1)\}$ for $c = 0$, $m \geq 2$ and $n \geq 6$, or if

$$a \leq 2\min\left[1, \left\{1 + \frac{(n-1)(b-c)}{(n+2)b}\right\}\frac{n-2}{m+1}c\right],$$

for $b \geq c > 0$, $m \geq 2$ and $n \geq 3$. This result for the squared-error loss is presented by Kubokawa (1987) when $d = 0$.

EXAMPLE 2.2. Setting $\phi = \max[(a-1)S_1/\{bS_2 + c(\bar{X} - \bar{Y})^2\}, 1]$ yields

$$\hat{\mu}_2(a, b, c) = \bar{X} + \min\left\{1, \frac{aS_1}{S_1 + bS_2 + c(\bar{X} - \bar{Y})^2}\right\}(\bar{Y} - \bar{X}),$$

which is minimax if the same conditions hold as in Example 2.1.

EXAMPLE 2.3. Define ϕ to be $\max[\{bS_2 + c(\bar{X} - \bar{Y})^2\}/S_1, 1]$. This gives the estimator

$$\hat{\mu}_3(a, b, c) = \bar{X} + \min\left[\frac{aS_1^2}{S_1^2 + \{bS_2 + c(\bar{X} - \bar{Y})^2\}^2}, \frac{aS_1}{S_1 + bS_2 + c(\bar{X} - \bar{Y})^2}\right] \times (\bar{Y} - \bar{X})$$

and satisfies conditions (b) and (c) of Theorem 2.1 for $t = 1$. Thus, from conditions (a), (d) and Lemma 2.2, $\hat{\mu}_3(a, b, c)$ is minimax, either if $a \leq 2\min\{1, b(n-7)/(m+1)\}$ for $c = 0$, $m \geq 2$ and $n \geq 8$, or if

$$a \leq 2\min\left[1, \left\{1 + \frac{(n-1)(b-c)}{(n+2)b}\right\}\frac{n-4}{m+3}c\right],$$

for $b \geq c > 0$; $m \geq 2$ and $n \geq 5$.

3. Admissible minimax estimators.

3.1. *Bayes equivariant estimators.* It is of great interest to find an admissible minimax estimator. In this section, we get Bayes equivariant estimators of μ relative to the loss (1.1) and apply Theorem 2.1 and a result of Brown (1966) to show their minimaxity and admissibility.

The problem defined in Section 1 remains invariant under the translation group whose transforms are

$$X_i \to X_i + c, \qquad Y_j \to Y_j + c, \qquad (\mu, \sigma_1^2, \sigma_2^2) \to (\mu + c, \sigma_1^2, \sigma_2^2), \qquad \hat{\mu} \to \hat{\mu} + c,$$

for $i = 1, \ldots, m$, $j = 1, \ldots, n$ and $-\infty < c < \infty$. Then (1.1) is quadratic so that we need only consider estimators which are functions of the minimal sufficient statistic $(\bar{X}, \bar{Y}, S_1, S_2)$. For simplicity, let $Z = \bar{Y} - \bar{X}$, $\eta = n/\sigma_2^2$ and $\rho = m\sigma_2^2/(n\sigma_1^2)$. It is easily seen that any estimator, equivariant under the preceding translation group, can be represented as $\bar{X} + \psi(Z, S_1, S_2)$ and that its risk function takes the form $R(\eta, \rho; \psi) = (1/m)E[1 + \rho\eta\psi^2 - 2\{\rho\eta/(1+\rho)\}Z\psi]$ since $E[\bar{X} - \mu|Z] = -Z/(1+\rho)$. If the nuisance parameter (η, ρ) has a prior distribution $H(\eta, \rho)$, the expected risk of an equivariant estimator $\bar{X} + \psi(Z, S_1, S_2)$ is $R(H; \psi) = \int\int R(\eta, \rho; \psi)H(d\eta, d\rho)$. We shall call the equivariant estimator which minimizes this Bayes risk as Bayes equivariant and denote it by $\hat{\mu}_H = \bar{X} + \psi_H(Z, S_1, S_2)$ in the present paper. We note that the Bayes risk $R(H; \psi_H)$ is finite since $R(H; \psi_H) \le R(H; 0) < \infty$. It is also remarked that the Bayes equivariant estimator is not necessarily Bayes in the usual sense. To obtain this Bayes equivariant estimator $\hat{\mu}_H$, we utilize the Fubini theorem and get

$$R(H; \psi_H) = \frac{1}{m}E\left[1 + E_H[\rho\eta|Z, S_1, S_2]\psi_H^2 - E_H\left[\frac{2\rho\eta}{1+\rho}\middle|Z, S_1, S_2\right]Z\psi_H\right], \tag{3.1}$$

where $E_H[\cdot|Z, S_1, S_2]$ designates the posterior expectation of the term in brackets, given (Z, S_1, S_2). The function minimizing (3.1) is $\psi_H = E_H[\rho\eta/(1+\rho)|Z, S_1, S_2]Z/E_H[\rho\eta|Z, S_1, S_2]$. Especially, take

$$H(d\eta, d\rho) \propto \frac{\rho^{n/2+\alpha+\varepsilon-1}}{(1+\rho)^{(m+n-1)/2+\varepsilon}}\eta^{\varepsilon-1/2}e^{-(\lambda/2)\eta}\,d\eta\,d\rho, \qquad \eta > 0,\ \rho > 0, \tag{3.2}$$

for some $\lambda > 0$, $\varepsilon > -1/2$ and $0 < \alpha < (m-1)/2$, which guarantee $\int\int H(d\eta, d\rho)$ is finite. Then, integrating out over η and making the transformation $\Theta = S_1\rho/S_2$ yields

$$\psi_H = Z\left\{1 + (S_2/S_1)\phi_H(S_1, S_2, Z^2)\right\}^{-1}, \tag{3.3}$$

where

$$\phi_H(S_1, S_2, Z^2) = \frac{\int_0^\infty \Theta^\alpha(\Theta/r)^{(m+n)/2+\varepsilon+1}\,d\Theta}{\int_0^\infty \Theta^{\alpha-1}(\Theta/r)^{(m+n)/2+\varepsilon+1}\,d\Theta} \tag{3.4}$$

and

$$r = S_2^2\Theta^2 + (S_1 + S_2 + Z^2 + \lambda)S_2\Theta + S_1(S_2 + \lambda).$$

Thus, the Bayes equivariant estimator $\hat{\mu}_H$ against $H(d\eta, d\rho)$ in (3.2) relative to the loss (1.1) is given by

$$\hat{\mu}_H = \bar{X} + \frac{1}{1 + (S_2/S_1)\phi_H\left(S_1, S_2, (\bar{X} - \bar{Y})^2\right)}(\bar{Y} - \bar{X}). \tag{3.5}$$

The estimator $\hat{\mu}_H$ is computationally complicated since it is defined by (3.4). Putting $a = b = 1$, $c = 0$ and $\phi = \phi_H$ in (2.1), we have $\hat{\mu}_{\phi_H}(1,1,0) = \hat{\mu}_H$, which shows that $\hat{\mu}_H$ belongs to the class of estimators of the form (2.1). So $\hat{\mu}_H$ is unbiased.

3.2. *Minimaxity and admissibility of the Bayes equivariant estimator.* We now show the minimaxity and admissibility of the Bayes equivariant estimator $\hat{\mu}_H$ under the loss (1.1). For this, we need the following lemma:

LEMMA 3.1. *For* $0 < \alpha < (m + n)/2 + \varepsilon$, $\phi_H(S_1, S_2, Z^2)$ *given by* (3.4) *is nondecreasing in* S_1 *and is nonincreasing in* S_2. *Moreover,*

$$\phi_H(S_1, S_2, Z^2) \geq \frac{\alpha}{(m + n)/2 + \varepsilon - \alpha}. \tag{3.6}$$

PROOF. We first prove that ϕ_H is nondecreasing in S_1. By differentiating ϕ_H with respect to S_1, we see that it is sufficient to show that

$$\frac{\int_0^\infty \Theta^{\alpha-1+N} r^{-N-1}(dr/dS_1)\, d\Theta}{\int_0^\infty \Theta^{\alpha+N} r^{-N-1}(dr/dS_1)\, d\Theta} \geq \frac{\int_0^\infty \Theta^{\alpha-1+N} r^{-N}\, d\Theta}{\int_0^\infty \Theta^{\alpha+N} r^{-N}\, d\Theta}, \tag{3.7}$$

where $N = (m + n)/2 + \varepsilon + 1$. Differentiating $(1/r)(dr/dS_1)$ with respect to Θ gives that it is nonincreasing in Θ. On the other hand, $1/\Theta$ is decreasing, so that (3.7) follows from Lemma 2.1. Hence, ϕ_H is nondecreasing in S_1. Similarly, it can be shown that ϕ_H is nonincreasing in S_2. Using properties of ϕ_H, we also have $\phi_H(S_1, S_2, Z^2) \geq \phi_H(0, \infty, Z^2)$. Making the transformation $x = \Theta/(1 + \Theta)$ and integrating over x gives $\phi_H(0, \infty, Z^2) = B(\alpha + 1, (m + n)/2 + \varepsilon - \alpha)/B(\alpha, (m + n)/2 + \varepsilon + 1 - \alpha) = \alpha/\{(m + n)/2 + \varepsilon + 1 - \alpha\}$. Thus, Lemma 3.1 is completely valid. □

On the basis of Lemma 3.1, we get

THEOREM 3.1. (i) *The Bayes equivariant estimator* $\hat{\mu}_H$ *given by* (3.5) *is minimax relative to the loss* (1.1) *provided* $m \geq 2$, $n \geq 6$, $\lambda > 0$, $\varepsilon > -1/2$ *and* $\alpha(m, n, \varepsilon) \leq \alpha < (m - 1)/2$, *where* $\alpha(m, n, \varepsilon) = (m + n + 2\varepsilon)(m + 1)/(2m + 4n - 18)$.

(ii) *The estimator* $\hat{\mu}_H$ *is admissible provided* $m \geq 4$, $n \geq 8$, $\lambda > 0$, $\varepsilon > 0$ *and* $0 < \alpha < (m - 3)/2$.

(iii) *The estimator* $\hat{\mu}_H$ *is admissible minimax provided* $m \geq 4$, $n \geq 8$, $\lambda > 0$, $\varepsilon > 0$ *and* $\alpha(m, n, \varepsilon) \leq \alpha < (m - 3)/2$.

REMARK 3.1. We can choose α satisfying condition (iii) in Theorem 3.1 if there is some $\varepsilon > 0$ such that $(m - 7)(n - 13) - 64 > 2(m + 1)\varepsilon$, which is possible whenever $(m - 7)(n - 13) > 64$. This shows that m and n need to be large in order for Theorem 3.1 to establish that $\hat{\mu}_H$ is admissible minimax.

PROOF OF THEOREM 3.1. To prove (i), it is only necessary to verify the conditions of Theorem 2.1 with $t = 0$. Conditions (a) and (c) are trivial. Condition (b) is evident from Lemma 3.1 for $0 < \alpha < (m + n)/2 + \varepsilon$. From Lemma 2.2(i) and (3.6), condition (d) is satisfied if $\{(m + n)/2 + \varepsilon - \alpha\}/\alpha \leq 2(n - 5)/(m + 1)$, which is equivalent to $\alpha(m, n, \varepsilon) \leq \alpha$. Combining the fact that $\hat{\mu}_H$ is Bayes equivariant for $\lambda > 0$, $\varepsilon > -1/2$ and $0 < \alpha < (m - 1)/2$, we get the conclusion.

To prove (ii), a result from Brown (1966) will be used. Let $z = \bar{y} - \bar{x}$, and let $f(\bar{x} - \mu, z|\eta, \rho)$ and $g(s_1, s_2|\eta, \rho)$ be joint densities of $(\bar{X} - \mu, Z)$ and (S_1, S_2), respectively. We first define a probability density function $p(\bar{x} - \mu, z, s_1, s_2)$ with a location parameter μ by

$$p(\bar{x} - \mu, z, s_1, s_2) = C\int\int \rho\eta f(\bar{x} - \mu, z|\eta, \rho) g(s_1, s_2|\eta, \rho) H(d\eta, d\rho), \tag{3.8}$$

where C is a normalization constant. The above integral is seen to be finite for $\lambda > 0$, $\varepsilon > -3/2$ and $0 < \alpha < (m - 3)/2$. Then from a result of Brown and Fox (1974), page 809, it is sufficient to prove the admissibility of $\hat{\mu}_H$ in the problem defined by (3.8) and the loss $(\hat{\mu} - \mu)^2$. Further for $\lambda > 0$, $\varepsilon > -1/2$ and $0 < \alpha < (m - 3)/2$, applying Theorem 2.1.1, Lemma 2.2.2 and Lemma 2.3.3 of Brown (1966), we can see that $\hat{\mu}_H$ is admissible if

$$\int\int\int\int |x|(x + \psi_H)^2 p(x, z, s_1, s_2)\, dx\, dz\, ds_1\, ds_2 < \infty, \tag{3.9}$$

$$\int\int\int\int (x^2 + \psi_H^2)|x + \psi_H| p(x, z, s_1, s_2)\, dx\, dz\, ds_1\, ds_2 < \infty, \tag{3.10}$$

where ψ_H is given by (3.3). Hence, (3.9) and (3.10) must be verified. The Fubini theorem first gives that (3.9) is equivalent to $\int\int Q(\eta, \rho; \psi_H) H(d\eta, d\rho) < \infty$, where

$$Q(\eta, \rho; \psi_H) = \frac{\rho\eta}{m}\int\int\int\int |x|(x + \psi_H)^2 f(x, z|\eta, \rho) g(s_1, s_2|\eta, \rho)\, dx\, dz\, ds_1\, ds_2.$$

Making the transformations

$$u = \sqrt{\eta}\sqrt{1 + \rho}\,\{x + z/(1 + \rho)\} \quad \text{and} \quad v = \sqrt{\eta}\sqrt{\rho/(1 + \rho)}\, z$$

and using (3.6), we have

$$\begin{aligned} Q(\eta, \rho; \psi_H) &= \frac{\rho}{m\sqrt{\eta}\sqrt{1 + \rho}\,(1 + \rho)} \\ &\quad \times \int\int\int\int \left| u - \frac{1}{\sqrt{\rho}} v \right| \left[u + \left\{ \frac{s_1(1 + \rho)}{s_1 + s_2\phi_H} - 1 \right\} \frac{1}{\sqrt{\rho}} v \right]^2 \\ &\qquad \times \frac{1}{2\pi}\exp\{-\tfrac{1}{2}(u^2 + v^2)\} g(s_1, s_2|\eta, \rho)\, du\, dv\, ds_1\, ds_2 \\ &\le \frac{2}{m\sqrt{\eta}\sqrt{1 + p}} \int\int \left(|u| + \frac{1}{\sqrt{\rho}}|v| \right) \\ &\qquad \times \left\{ u^2 + 2E\left[\left(1 + \frac{1}{C_0 W} \right)^2 + 1 \right] \frac{1}{\rho} v^2 \right\} \\ &\qquad \times \frac{1}{2\pi}\exp\{-\tfrac{1}{2}(u^2 + v^2)\}\, du\, dv, \end{aligned} \tag{3.11}$$

where $W = S_2/(\rho S_1)$ and $C_0 = \alpha/\{(m + n)/2 + \varepsilon - \alpha\}$. The r.h.s. of the inequality in (3.11) is integrable with respect to $H(d\eta, d\rho)$, provided $m \geq 4$, $n \geq 6$, $\lambda > 0$, $\varepsilon > 0$ and $0 < \alpha < (m-3)/2$. Thus, (3.9) holds. It is similarly shown that (3.10) is satisfied for $m \geq 4$, $n \geq 8$, $\lambda > 0$, $\varepsilon > 0$ and $0 < \alpha < (m-3)/2$. Therefore, part (ii) of the theorem is proved.

Part (iii) of the theorem is obtained by combining results (i) and (ii), which completes the proof of Theorem 3.1. □

REMARK 3.2. It is noticed in the proof of Theorem 3.1(i) that the minimaxity of $\hat{\mu}_H$ is guaranteed by a weaker condition $\alpha(m, n, \varepsilon) \leq \alpha < (m+n)/2 + \varepsilon$ unless it is necessary to be Bayes equivariant. It is also noted that Theorem 3.1(i) means that the estimator $\hat{\mu}_H$ is admissible minimax within the class of equivariant estimators.

REMARK 3.3. An interesting problem is to find admissible and Bayes equivariant estimators that beat both $\overline{X}$ and $\overline{Y}$ simultaneously. To obtain such estimators based on Theorem 2.1, we should consider a family of estimators of the form $\hat{\mu}_\phi(1,1,0)$, the special type of (2.1), which are better than both $\overline{X}$ and $\overline{Y}$. Then from Theorem 2.1 and considerations of symmetry, it follows that $\hat{\mu}_\phi(1,1,0)$ is better than both $\overline{X}$ and $\overline{Y}$ if $\phi(S_1, S_2, (\overline{X}-\overline{Y})^2)$ is nondecreasing in S_1 and nonincreasing in S_2, and if

$$(m+1)/\{2(n-5)\} \leq \phi\big(S_1, S_2, (\overline{X}-\overline{Y})^2\big) \leq 2(m-5)/(n+1),$$

for $m > 5$ and $n > 5$. The question is: Can $\hat{\mu}_H$, given by (3.5), belong to this family? Unfortunately, the answer is no because ϕ_H is not bounded above. In fact, making the transformation $\rho = S_2\Theta/S_1$ yields $\phi_H = (S_1/S_2)h(S_1, S_2, Z^2)$, where

$$h\big(S_1, S_2, Z^2\big) = \frac{\int_0^\infty \rho^\alpha \big(\rho/r^*\big)^{(m+n)/2+\varepsilon+1} d\rho}{\int_0^\infty \rho^{\alpha-1} \big(\rho/r^*\big)^{(m+n)/2+\varepsilon+1} d\rho}$$

and

$$r^* = S_1\rho^2 + \big(S_1 + S_2 + Z^2 + \lambda\big)\rho + S_2 + \lambda.$$

Further for $0 < \alpha < (m+n)/2 + \varepsilon$, the same argument as in the proof of Lemma 3.1 gives that $h(S_1, S_2, Z^2)$ is nonincreasing in S_1 and nondecreasing in S_2, so that $h(S_1, S_2, Z^2) \geq C_0$ for $C_0 = \alpha/\{(m+n)/2 + \varepsilon - \alpha\}$. Hence, we get the inequality $\phi_H \geq C_0 S_1/S_2$, which implies that ϕ_H is not bounded above. To obtain admissible estimators better than $\overline{X}$ and $\overline{Y}$ by use of our method, we would need at least to consider a different prior distribution on the nuisance parameter (σ_1^2, σ_2^2).

Acknowledgments. The author is extremely grateful to Professor N. Sugiura for his helpful criticism and valuable advice during the progress of this work. The author is also grateful to the Associate Editor and the referee for very valuable comments and suggestions.

REFERENCES

BHATTACHARYA, C. G. (1978). Yates type estimator of a common mean. *Ann. Inst. Statist. Math.* **30** 407–414.

BHATTACHARYA, C. G. (1980). Estimation of a common mean and recovery of interblock information. *Ann. Statist.* **8** 205–211.

BHATTACHARYA, C. G. (1984). Two inequalities with an application. *Ann. Inst. Statist. Math.* **36** 129–134.

BROWN, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* **37** 1087–1136.

BROWN, L. D. and COHEN, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. *Ann. Statist.* **2** 963–976.

BROWN, L. D. and FOX, M. (1974). Admissibility in statistical problems involving a location or scale parameter. *Ann. Statist.* **2** 807–814.

COHEN, A. and SACKROWITZ, H. B. (1974). On estimating the common mean of two normal distributions. *Ann. Statist.* **2** 1274–1282.

EXTON, H. (1978). *Handbook of Hypergeometric Integrals.* Wiley, New York.

GRAYBILL, F. A. and DEAL, R. B. (1959). Combining unbiased estimators. *Biometrics* **15** 543–550.

KHATRI, C. G. and SHAH, K. R. (1974). Estimation of location parameters from two linear models under normality. *Comm. Statist. A—Theory Methods* **3** 647–663.

KUBOKAWA, T. (1987). Estimation of a common mean of two normal distributions. *Tsukuba J. Math.* **11**, to appear.

SINHA, B. K. and MOUQADEM, O. (1982). Estimation of the common mean of two univariate normal populations. *Comm. Statist. A—Theory Methods* **11** 1603–1614.

ZACKS, S. (1970). Bayes and fiducial equivariant estimators of the common mean of two normal distributions. *Ann. Math. Statist.* **41** 59–69.

INSTITUTE OF MATHEMATICS
UNIVERSITY OF TSUKUBA
SAKURA-MURA NIIHARIGUN IBARAKI 305
JAPAN