The Annals of Statistics
1974, Vol. 2, No. 6, 1274–1282

ON ESTIMATING THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS

BY ARTHUR COHEN[1] AND HAROLD B. SACKROWITZ

Rutgers University

Consider the problem of estimating the common mean of two normal distributions. Two new unbiased estimators of the common mean are offered for the equal sample size case. Both are better than the sample mean based on one population for sample sizes of 5 or more. A slight modification of one of the estimators is better than either sample mean simultaneously for sample sizes of 10 or more. This same estimator has desirable large sample properties and an explicit simple upper bound is given for its variance. A final result is concerned with confidence estimation. Suppose the variance of the first population, say, is known. Then if the sample mean of that population, plus and minus a constant, is used as a confidence interval, it is shown that an improved confidence interval can be found provided the sample sizes are at least 3.

1. Introduction and summary. Consider random samples of size n from each of two independent normal distributions. The first distribution has mean θ and variance σ_x^2 and the second has mean θ and variance σ_y^2. Let $X' = (X_1, X_2, \cdots, X_n)$ and $Y' = (Y_1, Y_2, \cdots, Y_n)$ denote these samples. The problem is to estimate the common mean θ when the loss function is $(t - \theta)^2/\sigma_x^2$. This loss function is chosen for convenience. Squared error loss or squared error divided by a positive function of (σ_x^2, σ_y^2) could also be taken. This problem of estimating the common mean and the related problem of recovery of interblock information has been studied in several papers. For a brief bibliography and justification of some of the results studied here the reader is referred to the introduction of Brown and Cohen [2].

In this paper two new unbiased estimators for the common mean are suggested for the equal sample size case. Each estimator is uniformly better than the sample mean based on only one of the populations for $n \geqq 5$. A slight modification of one of the estimators is better than either sample mean for $n \geqq 10$. This is in contrast to the estimator studied by Graybill and Deal [3], which has such a property if and only if $n \geqq 11$. For this same new estimator, a very simple expression is derived which represents a bound on its risk.

One final result is concerned with confidence estimation. Suppose the variance of one of the populations, say σ_x^2, is known. Then if the sample mean of that population, plus and minus a constant, is used as a confidence interval, it is shown that an improved confidence interval can be found if $n \geqq 3$. In Sections

Received November 1972; revised November 1973.

[1] This author's research supported by NSF Grant No. GP-36566X1.

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

Key words and phrases. Common mean, unbiased estimators, minimax estimators, confidence intervals, inter-block information.

2 and 3 the new estimators are given and discussed. The confidence interval result is given in Section 4.

2. New unbiased estimator of the common mean. In this section we find an estimator which is unbiased and minimax for $n \geqq 5$, and we also compute its risk. The given estimator is not based only on a sufficient statistic. We are able to "Rao–Blackwellize" it and thus get an estimator which is better. The resulting estimator is thus unbiased and minimax. We have a bound for its risk. Furthermore, it has sensible monotonicity properties and has desirable large sample properties.

Define $\bar{X} = \sum_{i=1}^n X_i/n$, $\bar{Y} = \sum_{i=1}^n Y_i/n$, $S_x^2 = \sum_{i=1}^n (X_i - \bar{X})^2$, $S_y^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2$, $z = S_y^2/S_x^2$, $\tau = \sigma_y^2/\sigma_x^2$. Let a_i, $i = 1, 2, \cdots, n-1$ be $n \times 1$ vectors which are a set of orthonormal contrasts. Note that the $2n \times 1$ vectors $[a_i', 0]$, $i = 1, 2, \cdots, n-r-1$, $[a_i', a_i']$, $i = n-r, \cdots, n-1$, $[a_i', -a_i']$, $i = n-r, \cdots, n-1$, are all orthogonal contrasts. Define $U_i = a_i'X$, $i = 1, 2, \cdots, n-r-1$; $V_i = [a_i', a_i'][{}^X_Y]$, $i = n-r, \cdots, n-1$. $W_i = [a_i', -a_i'][{}^X_Y]$, $i = n-r, \cdots, n-1$. Clearly U_i, V_i, W_i are independent normal random variables with zero means and variances σ_x^2, $(\sigma_x^2 + \sigma_y^2)$ and $(\sigma_x^2 + \sigma_y^2)$ respectively. Furthermore, U_i, V_i, W_i are independent of $\bar{X}$ and $\bar{Y}$. Finally let $S^2_{n-r-1} = \sum_{i=1}^{n-r-1} U_i^2$, and $S^2 = \sum_{i=n-r}^{n-1} (V_i^2 + W_i^2)$. Clearly S^2_{n-r-1} is distributed as a $\sigma_x^2\chi^2(n-r-1)$ and S^2 is distributed independently of S^2_{n-r-1}, as a $(\sigma_x^2 + \sigma_y^2)\chi^2(2r)$.

Consider estimators of the form

$$(1 - CS^2_{n-r-1}/S^2)\bar{X} + C(S^2_{n-r-1}/S^2)\bar{Y}\,. \tag{2.1}$$

Clearly the estimators in (2.1) are unbiased. It is easy to show that the best value of C in (2.1) is $C_r = 2(r-2)/(n-r+1)$ and that (2.1) is better than $\bar{X}$ for all C such that $0 < C < 2C_r$. In fact for $C = C_r$ the estimator in (2.1) has risk

$$(1/n)[1 - [1/(1+\tau)](n-r-1)(r-2)/(n-r+1)(r-1)]\,. \tag{2.2}$$

Thus for any value of r ranging from 3 to $(n-2)$ the estimator in (2.1) with $C = C_r$ is unbiased and better than $\bar{X}$. For any given n we seek the optimal value of r and hence of C_r by minimizing (2.2). We find that (2.2) is minimized by choosing $r = (n+1)/2$ for n odd and $r = n/2$ for n even. For these values of r we come up with the following estimators:

$$[1 - 2[(n-3)/(n+1)](S^2_{(n-3)/2}/S^2)]\bar{X} + 2[(n-3)/(n+1)](S^2_{(n-3)/2}/S^2)\bar{Y} \tag{2.3}$$

for n odd;

$$[1 - 2[(n-4)/(n+2)](S^2_{n/2}/S^2)]\bar{X} + 2[(n-4)/(n+2)](S^2_{n/2}/S^2)\bar{Y} \tag{2.4}$$

for n even.

The risk for (2.3) is

$$(1/n)[1 - (1/(1+\tau))((n-3)^2/(n-1)(n+1))]\,. \tag{2.5}$$

The risk for (2.4) is

$$(1/n)[1 - (1/(1+\tau))((n-4)/(n+2))]\,. \tag{2.6}$$

From (2.5) and (2.6) it is clear that improvement over $\bar{X}$ can be made provided $n \geqq 5$.

Now note that a set of sufficient statistics for this problem is $(\bar{X}, \bar{Y}, S_x^2, S_y^2)$. Since the estimators in (2.3) and (2.4) are not based on the set of sufficient statistics it follows by a version of the Rao–Blackwell Theorem that the conditional expectations of (2.3) and (2.4), given $(\bar{X}, \bar{Y}, S_x^2, S_y^2)$, are better estimators than (2.3) or (2.4) and hence, than $\bar{X}$.

We derive the estimator which results from taking the conditional expectation of (2.3) or (2.4). We prove

LEMMA 2.1. *The unique unbiased estimate of* $(1/(1+\tau))$ *based on* (S_x^2, S_y^2) *is*

$$\begin{aligned} G(S_x^2, S_y^2) = G^*(z) &= F(1, (3-n)/2, (n-1)/2; z) \qquad \textit{for } 0 \leqq z \leqq 1 \\ &= [(n-3)/(n-1)](1/z)F(1, (5-n)/2, (n+1)/2; 1/z) \\ &\qquad \textit{for } z \geqq 1\,, \end{aligned} \tag{2.7}$$

where F is the hypergeometric function. (Often this hypergeometric function is denoted by ${}_2F_1$.)

PROOF. The method of proof is essentially the same as used by Olkin and Pratt [4].

We are now ready to prove

THEOREM 2.1. *The estimators*

$$[1 - C_n^* G^*(z)]\bar{X} + C_n^* G^*(z)\bar{Y}\,, \tag{2.8}$$

where $C_n^* = (n-3)^2/(n+1)(n-1)$, *for n odd, and*

$$C_n^* = (n-4)/(n+2)\,, \qquad \textit{for} \quad n \quad \textit{even}, \tag{2.9}$$

are unbiased and minimax for all $n \geqq 5$. *Furthermore, these estimators are better than the estimators in* (2.3) *and* (2.4) *and their risks are bounded by* (2.5) *and* (2.6) *respectively.*

PROOF. The proof is accomplished by showing that the conditional expectation of (2.3) and (2.4), given the set of sufficient statistics, is (2.8). An application of Lemma 2.1 is used to complete the proof of this theorem.

To facilitate computation of the estimator in (2.8) tables can be obtained from tables in Olkin and Pratt [4]. That is, from (2.7) and Olkin and Pratt's (3.3) and (3.5) it can be shown that

$$\begin{aligned} G^*(z) &= (1/2) + H(r')/2 \qquad \text{for} \quad 0 \leqq z \leqq 1 \\ G^*(z) &= (1/2) - H(r')/2 \qquad \text{for} \quad z \geqq 1\,, \end{aligned} \tag{2.10}$$

where $z = (r'+1)/(1-r')$, our $n = N+1$, where N represents the degrees of freedom in Olkin and Pratt's Table 2, page 208, and $H(r')$ is defined on page 208.

We conclude this section with the following remarks:

REMARK 2.1. For $n = \infty$, it is easy to verify, using (3.9) of Olkin and Pratt

[4], that $G^*(z) = 1/(1+z)$. Since $C_n^* \to 1$ as $n \to \infty$, we see that the weights given to the sample means in (2.8), are converging strongly to the optimal weights in the case where the variances are known.

REMARK 2.2. Recall that the estimator (2.1) is better than $\bar{X}$ for all C such that $0 < C < 2C_r$. This fact, the derivation of (2.8), and symmetry considerations imply that

$$[1 - G^*(z)]\bar{X} + G^*(z)\bar{Y} \tag{2.11}$$

is better than both $\bar{X}$ and $\bar{Y}$ simultaneously if $2C_n^* \geqq 1$, where C_n^* is given in (2.9). We see that this is true for $n \geqq 10$.

REMARK 2.3. If the sample sizes in each population are unequal, modifications of the ideas in this section could be made to derive similar type results. We do not pursue this matter here.

REMARK 2.4. How does the estimator (2.8) compare with other estimators already proposed for the common mean? For large values of n, the estimator (2.8) is essentially the same as that proposed by Graybill and Deal [3], namely

$$[1 - 1/(1+z)]\bar{X} + [1/(1+z)]\bar{Y}\,. \tag{2.12}$$

For moderate values of n, say above 5, the estimators are not quite comparable to (2.12). That is, (2.8) is better than $\bar{X}$, for $n \geqq 5$. The estimator (2.12) is not better than either $\bar{X}$ or $\bar{Y}$ for $n < 11$. Hence the risks of these estimators are non-comparable. The reason is, that $G^*(z)$ is being multiplied by a constant less than one. If one compares (2.11) with (2.12), (2.11) would be preferred for $n = 10$. The function $G^*(z) \geqq 1/(1+z)$ for $0 \leqq z \leqq 1$, and $G^*(z) \geqq 1/(1+z)$ for $z > 1$. Since $G^*(z)$ is the weight given to $\bar{Y}$, and (2.11) is preferable to (2.12) for $n = 10$, it appears that the risk of (2.11) would be less than the risk for (2.12) for small and large values of τ when n ranges from 5 upward. The associate editor noted that the risk for $\tau > 1$ for (2.12) is

$$\begin{aligned}[(n-1)(n+1)(1+\tau)/4n]\, &\textstyle\sum_{n=0}^{\infty} \frac{g(m \mid \omega, (n-1)/2)}{(n-1+m)(n+m)} \\ &+ (1/n) \textstyle\sum_{m=0}^{\infty} \frac{g(m \mid \omega, (n-1)/2)m}{n-1+m}\,,\end{aligned} \tag{2.13}$$

where

$$\begin{aligned} g(m \mid \omega, (n-1)/2) \\ = \{\Gamma([(n-1)/2] + m)/\Gamma((n-1)/2)\Gamma(m+1)\}\omega^m(1-\omega)^{(n-1)/2}\,, \end{aligned}$$

$m = 0, 1, 2, \cdots$ and $\omega = (1 - 1/\tau)$.

The method used to compute (2.13) could be used to compute the risk of the estimator in Theorem 2.2 of Brown and Cohen [2]. This estimator, for values of $n \geqq 6$, would be comparable to the estimator in (2.8). For $n = 5$, (2.8) is preferable to the Brown–Cohen counterpart. The same assessment comparing

(2.11) and (2.12) would apply to the comparison of (2.8) and the Brown–Cohen counterpart. For small values of n, perhaps the most reasonable estimators are those offered in Brown and Cohen, Theorem 2.1. However, this too does not appear to be easily established.

3. Other unbiased estimators better than the sample mean. In this section, for samples of size 5 or more, unbiased estimators are given which are uniformly better than the sample mean based on only one population.

The problem and model are the same as in the previous section.

Now consider the estimator

$$(3.1)\qquad (X_1 + X_2)/n + (1 - [(n-4)/3][(X_1 - X_2)^2/2\|X^{(2)} - Y^{(2)}\|^2]) \times (\bar{X}^{(2)} - \bar{Y}^{(2)}) + \bar{Y}^{(2)},$$

where $X^{(2)\prime} = (X_3, X_4, \cdots, X_n)$, $\bar{X}^{(2)} = \sum_{i=3}^n X_i/n$, $Y^{(2)}$, $\bar{Y}^{(2)}$ are defined similarly; and $\|X\|^2 = \sum_{i=1}^n X_i^2$. We prove

THEOREM 3.1. *The estimator in* (3.1) *is unbiased, and for $n \geqq 5$, it is uniformly better than $\bar{X}$.*

PROOF. Unbiasedness follows by observing that $E[\bar{Y}^{(2)} + (X_1 + X_2)/n] = \theta$, and that the distribution of $(\bar{X} - \bar{Y})$ is symmetric about zero. The risk of (3.1) is

$$\begin{aligned}
&(1/\sigma_x^2)E\{(X_1 + X_2)/n + (1 - [(n-4)/3][X_1 - X_2)^2/2\|X^{(2)} - Y^{(2)}\|^2] \\
&\qquad \times (\bar{X}^{(2)} - \bar{Y}^{(2)}) + \bar{Y}^{(2)} - (2\theta/n) - (n-2)\theta/n\}^2 \\
(3.2)\qquad &= (2/n^2) + [(n-2)/n^2\sigma_x^2]E\{(1 - [(n-4)/3] \\
&\qquad \times [(X_1 - X_2)^2/2\|X^{(2)} - Y^{(2)}\|^2]) \textstyle\sum_{i=3}^n (X_i - Y_i)/(n-2)^{\frac12} \\
&\qquad + \textstyle\sum_{i=3}^n (Y_i - \theta)/(n-2)^{\frac12}\}^2 .
\end{aligned}$$

We can now use results of Stein [5], pages 362–365. That is, in Stein's equations (3.3) and (3.4) use $Y^{(2)}$ for Z, $(n-2)$ for p, 1 for n, 1 for α, $(X_1 - X_2)^2/2$ for S, $X^{(2)}$ for Y, $\boldsymbol{\theta}$ for η, where $\boldsymbol{\theta}' = (\theta, 0, \cdots, 0)$, $1/(n-2)^{\frac12}$ for γ. Thus by the remark on the bottom of page 364 it follows that the expected value of the bracketed term in (3.2) is less than σ_x^2. This proves that the risk of (3.1) is less than the risk of $\bar{X}$ provided $(n-2) \geqq 3$, or $n \geqq 5$. This completes the proof of the theorem.

We next offer a corollary to Theorem 3.1. If we let $S = (X_1 - X_2)^2/2$, $F = \|X^{(2)} - Y^{(2)}\|^2/S$ and consider the class of estimators

$$(3.3)\qquad (X_1 + X_2)/n + (1 - r(F, S)/F)(\bar{X}^{(2)} - \bar{Y}^{(2)}) + \bar{Y}^{(2)}$$

we can prove

COROLLARY 3.1. *If* (i) *$r(F, S)$ is, for each fixed S, monotone non-decreasing in F,* (ii) *for each fixed F, monotone nonincreasing in S, and* (iii) $0 \leqq r \leqq 2(n-4)/3$, *then the estimator in* (3.3) *is unbiased and minimax.*

PROOF. The proof follows the same argument as given in Theorem 3.1. The final step, to show that the expectation of the analogue of the bracketed term in (3.2) is less than or equal to σ_x^2, is accomplished by referring to a theorem of Strawderman [6], page 3 and using the same reasoning as in the Stein paper.

We observe that the estimators in (3.3) can be improved on in two different ways. First, note that the risk of (3.3) can be rewritten as

$$\begin{aligned} & E\{(X_1 + X_2 - 2\theta)/n + (1 - r(F, S)/F)(\textstyle\sum_{i=3}^n (X_i - \theta)/n \\ (3.4) \qquad & \qquad + (r(F, S)/F)(\textstyle\sum_{i=3}^n (Y_i - \theta)/n)\}/\sigma_x^2 \\ & = 2/n^2 + [(n-2)/n^2\sigma_x^2]E\{(1 - r(F, S)/F)^2\sigma_x^2 + (r(F, S)/F)^2\sigma_y^2\}. \end{aligned}$$

It is clear from (3.4) that by replacing $r(F, S)/F$ with $\min(1, r(F, S)/F)$, we get a similar risk. To illustrate a second method of improvement we refer back to the estimate (3.1). The method would apply to (3.3) and even to the estimators above which improve on (3.3). Call (3.1) t_{12}. For $i < j$, let t_{ij} be the same as (3.1) except X_i replaces X_1, X_j replaces X_2, replace $X^{(2)}$ by $X^{(ij)}$, where $X^{(ij)}$ is the $(n-2)$ vector X with X_i and X_j missing. Similarly define $Y^{(ij)}$. Since t_{ij}, $i, j = 1, 2, \cdots, n$, $i < j$, are C_2^n estimators, all with the same risk it follows that the randomized estimator which chooses t_{ij} with probability $1/C_2^n$ has the same risk as any t_{ij}. Hence by a well-known theorem it follows that $\sum_{1 \leqq i < j \leqq n} t_{ij}/C_2^n$ is a better estimator than any t_{ij}.

4. Improved confidence interval. In this section we assume that σ_x^2 is known. For this case we show that the confidence interval $\bar{X} \pm h$, for h a positive constant, can be improved on, provided the sample sizes are at least 3. The improved confidence interval will have length equal to $2h$, and will have probability of coverage uniformly greater than $\alpha = [\Phi(n^{\frac{1}{2}}h) - \Phi(-n^{\frac{1}{2}}h)]$ where Φ is the cdf of the standard normal distribution.

To facilitate the proof of the result we need a pair of lemmas. In the first lemma we paraphrase a theorem of L. D. Brown [1] on inadmissibility of estimators for non-continuous loss functions. We apply Brown's theorem in the second lemma which is concerned with a loss which is a sum of non-continuous losses. The lemmas will be used in the proof of the result.

In order to state the first lemma we let Z be an $n \times 1$ random vector which is normally distributed with mean vector μ and covariance ρI_n where ρ is a known constant. Let $p(Z)$ denote the density of Z when $\mu = 0$, let p' denote the $1 \times n$ row vector whose ith component is the derivative of $p(z)$ with respect to Z_i, and let $p''_{ij}(Z)$ denote the second partial derivative of $p(Z)$ with respect to Z_i and then Z_j. Clearly, for the case here where p is the normal density, all such partial derivatives exist and are continuous. Now let $W(t, \mu) = W(t - \mu)$ represent the loss function for the problem of estimating μ. Hence $W(Z)$ would represent the loss if μ were estimated by Z and $\mu = 0$. Finally let R_0 denote the risk of the procedure where μ is estimated by Z. We are now ready to state

LEMMA 4.1. *Let $n \geqq 3$ and $R_0 < \infty$. Suppose*

$$(4.1) \qquad \text{(i)} \quad \textstyle\int W(Z)\|Z\|^4\|p'(Z)\|\,dZ < \infty\,;$$

(ii) *there is a* $\gamma > 0$ *and a* $C > 0$ *such that*

$$\int W(Z)||Z||^2|p''_{ij}(Z + G(Z)| \, dZ < C \tag{4.2}$$

for all $G(Z)$ *such that* $||G(Z)|| < \gamma$, $1 \leqq i, j \leqq n$;

(iii) *the* $n \times n$ *matrix*

$$M = \int W(Z)\{Zp'(Z)\} \, dZ + (\int W(Z)p(Z)I_n \, , \tag{4.3}$$

is nonsingular.

Define ε *by the* $n \times 1$ *vector point estimator*

$$\varepsilon(Z) = [(I + B)/(a + ||Z||^2)]Z \, , \tag{4.4}$$

where $B = (1/b)M^{-1}$. *Then there exist constants* a, b *such that the risk of the procedure* ε *is less than* R_0 *for all* μ.

PROOF. Brown [1], page 1132.

Now consider the region $[Z_i \pm h]$, $i = 1, 2, \cdots, n$, for $n \geqq 3$. Also let

$$W(t - \mu) = \sum_{i=1}^n (1 - I_{[t_i - h, t_i + h]}(\mu)) = \sum_{i=1}^n (1 - I_{[-h,h]}(t_i - \mu)) \tag{4.5}$$

where $I_{[a,b]}(\mu)$ is the indicator function. (Note that for this loss function we use the terminology "region" and "point estimate" interchangeably.) We prove

LEMMA 4.2. *Let* $n \geqq 3$ *and let* W *be as in* (4.5). *Then the region* $\varepsilon_i(Z) \pm h$, $i = 1, 2, \cdots, n$, *where* $\varepsilon(Z)$ *is defined in* (4.4), *is a better region than* $Z_i \pm h$, $i = 1, 2, \cdots, n$.

PROOF. The proof follows from Lemma 3.1 by verifying (4.1), (4.2), and showing that M, defined in (4.3), is nonsingular. We omit the details.

Now we return to the model of this section. That is, we observe a random sample of size n from a normal distribution with mean θ and known variance, which without loss of generality is taken to be 1. The sample is represented by the $n \times 1$ vector X which is said to be multivariate normal with mean vector $\boldsymbol{\theta}$ and covariance I_n. The vector $\boldsymbol{\theta}$ has all its components equal to θ. We also observe an $n \times 1$ random vector Y which is assumed to be multivariate normal with mean vector $\boldsymbol{\theta}$ and covariance matrix $\sigma_y^2 I_n$. Let $\bar{X} \pm h$ be a confidence interval for θ. The coverage probability for such an interval is $\alpha = [\Phi(n^{\frac{1}{2}}h) - \Phi(-n^{\frac{1}{2}}h)]$. We prove

THEOREM 4.1. *Let* $n \geqq 3$. *Then if* $0 < \alpha < 1$, *there exist constants* a^* *and* b^*, $0 < b^* < a^* < \infty$, *such that the confidence interval*

$$\left[\left(1 - \frac{b^*}{a^* + ||X - Y||^2}\right)(\bar{X} - \bar{Y}) + \bar{Y} \pm h\right], \tag{4.6}$$

has coverage probability greater than α, *for all* (θ, σ_y^2).

PROOF. Let P be an $n \times n$ orthogonal matrix whose first row has all its components equal to $1/n^{\frac{1}{2}}$. Let $V = (1/n^{\frac{1}{2}})PX$, so that V is multivariate normal with mean vector ν and covariance matrix I/n. Note that $\nu' = (\nu_1, \nu_2, \cdots, \nu_n) = (\theta, 0, \cdots, 0)$, and that V_1, the first component of V, is equal to $\bar{X}$. Now let $W = (1/n^{\frac{1}{2}})P\bar{Y}$ so that $W_1 = \bar{Y}$ and $EW_1 = \theta$, $EW_i = 0$, $i = 2, \cdots, n$. Suppose that the problem is to observe (V, W) and find a region for ν with loss function $\sum_{i=1}^n (1 - I_{[t_i \pm h]}(\nu_i))$. Consider the region

$$[V_i \pm h], \qquad i = 1, 2, \cdots, n \tag{4.7}$$

with risk $(n - n\alpha)$. Now consider a region of the form

$$\left(1 - \frac{b}{a + \|V - W\|^2}\right)(V_i - W_i) + W_i \pm h, \qquad i = 1, 2, \cdots, n. \tag{4.8}$$

For each fixed value $W = w$, there exists an a and b, independent of w, such that

$$\left(1 - \frac{b}{a + \|V - w\|^2}\right)(V_i - w_i) + w_i \pm h, \tag{4.9}$$

has risk less than $n - n\alpha$. This follows from Lemma 4.2 since $(V - W)$ given $W = w$, is multivariate normal with mean vector $\nu - w$ and convariance I/n. Since a and b are found independently of the mean vector in the above lemma, it follows that the region in (4.8) has risk smaller than $(n - n\alpha)$. But the risk of (4.8) is

$$\begin{aligned} &n - \textstyle\sum_{i=1}^n P\left\{\left|\left(1 - \frac{b}{a + \|V - W\|^2}\right)(V_i - W_i) + (W_i - \nu_i)\right| < h\right\} \\ &\quad = n\left(1 - P\left\{\left|\left(1 - \frac{b}{a + \|V - W\|^2}\right)(V_1 - W_1) + (W_1 - \nu_1)\right| < h\right\}\right). \end{aligned} \tag{4.10}$$

To get (4.10) we used the fact that the random variables

$$\left[\left(1 - \frac{b}{a + \|V - W\|^2}\right)(V_i - W_i) + (W_i - \nu_i)\right] \qquad i = 1, 2, \cdots, n$$

are identically distributed. Since the quantity on the right-hand side of (4.10) is less than $n(1 - \alpha)$ it follows that the confidence interval

$$\left[\left(1 - \frac{b}{a + \|V - W\|^2}\right)(V_1 - W_1) + W_1 \pm h\right] \tag{4.11}$$

has coverage probability greater than α. If we rewrite (4.11) in terms of X and Y, we get

$$\left[\left(1 - \frac{nb}{na + \|X - Y\|^2}\right)(\bar{X} - \bar{Y}) + W_1 \pm h\right]. \tag{4.12}$$

By letting $b^* = nb$ and $a^* = na$, (4.12) coincides with (4.6) and this completes the proof of the theorem.

Acknowledgment. The authors wish to thank the referee and associate editor for helpful remarks.

REFERENCES

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

[2] 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.

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

[4] OLKIN, I. and PRATT, J. (1958). Unbiased estimation of certain correlation coefficients. *Ann. Math. Statist.* **29** 201–211.

[5] STEIN, C. (1966). An approach to the recovery of inter-block information in balanced incomplete block designs. *Research Papers in Statistics* (Neyman Festschrift, F. N. David, ed.). 351–366. Wiley, New York.

[6] STRAWDERMAN, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. *Ann. Statist.* **1** 1189–1194.

STATISTICS CENTER
RUTGERS UNIVERSITY
NEW BRUNSWICK, N. J. 08903