Ann. Inst. Statist. Math.
Vol. 41, No. 3, 477-484 (1989)

CLOSER ESTIMATORS OF A COMMON MEAN IN THE SENSE OF PITMAN

TATSUYA KUBOKAWA

Institute of Mathematics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

(Received April 21, 1987; revised December 27, 1988)

Abstract. Consider the problem of estimating the common mean of two normal populations with different unknown variances. Suppose a random sample of size m is drawn from the first population and a random sample of size n is drawn from the second population. The paper gives a family of estimators closer than the sample mean of the first population in the sense of Pitman (1937, *Proc. Cambridge Phil. Soc.*, **33**, 212-222). In particular, the Graybill-Deal estimator (1959, *Biometrics*, **15**, 543-550) is shown to be closer than each of the sample means if $m \geq 5$ and $n \geq 5$.

Key words and phrases: Pitman closeness, common mean, Graybill-Deal estimator.

1. Introduction

Let $(X_1,\dots,X_m)$ and $(Y_1,\dots,Y_n)$ be independent random samples from two normal populations with a common unknown mean μ and unknown variances σ_1^2 and σ_2^2, respectively. Also, 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. Based on $\bar{X}$, $\bar{Y}$, S_1 and S_2, we want to estimate the common mean μ.

This problem of estimating the common mean and the related problem of recovery of interblock information have been studied in several papers. For a brief bibliography the reader is referred to Bhattacharya (1980). Graybill and Deal (1959) considered the combined estimator

$$\hat{\mu}_{\mathrm{GD}} = \left(\frac{m-1}{S_1} \bar{X} + \frac{n-1}{S_2} \bar{Y} \right) \bigg/ \left(\frac{m-1}{S_1} + \frac{n-1}{S_2} \right), \tag{1.1}$$

and showed that $\hat{\mu}_{\mathrm{GD}}$ has a smaller variance than both $\bar{X}$ and $\bar{Y}$ if and only if $m \geq 11$ and $n \geq 11$, which was, later, corrected by Khatri and Shah (1974) as $(m \geq 11, n \geq 11)$, $(m = 10, n \geq 19)$ or $(m \geq 19, n = 10)$. This means that the combined estimator does not always dominate the uncombined estimator for sample sizes smaller than 10. Intuitively, however, it seems that the combined estimator is superior to the uncombined one for smaller sample sizes, if we choose another criterion for comparing estimators.

The criterion we employ here is the Pitman closeness, which is defined as follows: For two estimators $\hat{\mu}_1$ and $\hat{\mu}_2$ of μ, $\hat{\mu}_1$ is said to be *closer* than $\hat{\mu}_2$ in the sense of Pitman (1937) if and only if

$$P\{(\hat{\mu}_1 - \mu)^2 \leq (\hat{\mu}_2 - \mu)^2\} \geq 1/2 ,$$

uniformly with respect to unknown parameters. The Pitman closeness was used by Sugiura (1984) for estimating the normal covariance matrix, and was discussed by Peddada and Khattree (1986) and Rao *et al.* (1986). Theorem 2.1 of Sen (1986) gives the condition on variance for one estimator being closer than another. If σ_1^2 and σ_2^2 are known, it follows from his theorem that the maximum likelihood estimator $(m\sigma_1^{-2}\bar{X} + n\sigma_2^{-2}\bar{Y})/(m\sigma_1^{-2} + n\sigma_2^{-2})$ is always closer than both $\bar{X}$ and $\bar{Y}$. However, since σ_1^2 and σ_2^2 are unknown, we cannot use his result.

In this paper, we obtain a family of estimators which are closer than $\bar{X}$ in the sense of Pitman, and present the interesting example that the Graybill-Deal estimator $\hat{\mu}_{GD}$ is closer than both $\bar{X}$ and $\bar{Y}$ if $m \geq 5$ and $n \geq 5$ as is shown in Example 2.1. This demonstrates that the Graybill-Deal estimator has a desirable property for smaller sample sizes.

For estimation of a mean vector of a p-variate normal distribution with unknown variance, Sen *et al.* (1989) recently showed that the James-Stein type estimator dominates the usual one relative to the Pitman closeness criterion if $p \geq 2$. This result can be proved based on the monotonicity of a probability with respect to the noncentrality parameter. For our purpose, the same argument as in the proof is useful. Our model, however, is different from the Stein problem, and the monotonicity with respect to the variance ratio σ_2^2/σ_1^2 is essential in our proof.

2. Main result

For nonnegative constants a, b and c, consider the estimators of the form

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

where $R = \{bS_2 + c(\bar{X} - \bar{Y})^2\}/S_1$ and ϕ is a positive valued function. These are unbiased estimators of μ, and were proposed by Kubokawa (1987*b*). Then we get

THEOREM 2.1. *Assume that*

$$\phi(S_1, S_2, (\bar{X} - \bar{Y})^2) \geq \frac{(m-1)a}{2(n-3)b}, \tag{2.2}$$

for $m \geq 2$, $n \geq 4$ and $0 < a \leq 4/3$. Then $\hat{\mu}_\phi(a,b,c)$ given by (2.1) *is closer than $\bar{X}$ in the sense of Pitman.*

When $a = 1$ and $c = 0$, in particular, the consideration of symmetry yields

COROLLARY 2.1. *When $a = 1$ and $c = 0$, assume that*

$$\frac{m-1}{2(n-3)} \leq b\phi(S_1, S_2, (\bar{X} - \bar{Y})^2) \leq \frac{2(m-3)}{n-1}, \tag{2.3}$$

for m, $n \geq 4$. Then $\hat{\mu}_\phi(1,b,0)$ is closer than both $\bar{X}$ and $\bar{Y}$ in the sense of Pitman.

Note. We can choose $b\phi(S_1, S_2, (\bar{X} - \bar{Y})^2)$ which satisfies the condition (2.3) if and only if

$$(3m - 11)(3n - 11) \geq 16,$$

which is equivalent to $(m = 4, n \geq 9)$, $(m \geq 5, n \geq 5)$ or $(m \geq 9, n = 4)$.

Some examples of closer estimators based on Theorem 2.1 are given below. They were discussed by Kubokawa (1987*b*) under a quadratic loss function.

Example 2.1. Define ϕ to be $1 + d/\{bS_2 + c(\bar{X} - \bar{Y})^2\}$ in (2.1) for $d \geq 0$. This gives the estimator

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

which includes the estimator $\hat{\mu}_1(1,(m-1)/(n-1),0,0)$ ($= \hat{\mu}_{GD}$) 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*a*). Then, Theorem 2.1 presents that $\hat{\mu}_1(a,b,c,d)$ is closer than $\bar{X}$ if

$$m \geq 2, \quad n \geq 4, \quad 0 < a \leq 4/3 \quad \text{and} \quad b \geq (m-1)a/\{2(n-3)\}.$$

For the second sample, this condition always requires the size $n \geq 4$, although, in the sense of minimizing variance, $n \geq 3$ is sufficient when $c > 0$. This fact results from neglecting the information about $(\bar{X} - \bar{Y})^2$ in the proof of Theorem 2.1. Corollary 2.1 also gives that Graybill-Deal estimator $\hat{\mu}_{GD}$ is closer than both $\bar{X}$ and $\bar{Y}$ if $m \geq 5$ and $n \geq 5$.

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 closer than $\bar{X}$ if the same condition as in Example 2.1 holds.

Example 2.3. Setting $\phi = \max\,[\{bS_2 + c(\bar{X} - \bar{Y})^2\}/S_1, 1]$ gives

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

which is closer than $\bar{X}$ for the same condition as in Example 2.1.

To prove Theorem 2.1, we need the following lemma.

LEMMA 2.1. *Let X be a positive random variable such that $E[X^{-1}]$ is finite. Then for* $0 \leq p \leq 1$,

$$\frac{1}{E[\{(1-p) + pX\}^{-1}]} \geq \min\left\{1, \frac{1}{E[X^{-1}]}\right\}.$$

PROOF. Observe that $\min\{1, 1/E[X^{-1}]\} \leq (1-p) + p/E[X^{-1}] = E[\{(1-p) + pX\}X^{-1}]/E[X^{-1}]$. Since $\{(1-p) + pX\}^{-1}$ and $\{(1-p) + pX\}X^{-1}$ are monotone in the same direction with respect to X,

$$\frac{1}{E[\{(1-p) + pX\}^{-1}]} \geq \frac{E[\{(1-p) + pX\}X^{-1}]}{E[X^{-1}]},$$

which establishes Lemma 2.1.

PROOF OF THEOREM 2.1. From the definition of the Pitman closeness, we shall prove that

$$P\{(\hat{\mu}_\phi(a,b,c)-\mu)^2 \le (\bar{X}-\mu)^2\} \ge 1/2\,, \tag{2.4}$$

uniformly. Note that $(\hat{\mu}_\phi(a,b,c)-\mu)^2 \le (\bar{X}-\mu)^2$ if and only if

$$\frac{a}{1+R\phi}(\bar{Y}-\bar{X})^2 + 2(\bar{X}-\mu)(\bar{Y}-\bar{X}) \le 0\,. \tag{2.5}$$

From the condition (2.2) and the fact that $R \ge bS_2/S_1$, the inequality (2.5) holds if

$$\frac{a}{1+AS_2/S_1}(\bar{Y}-\bar{X})^2 + 2(\bar{X}-\mu)(\bar{Y}-\bar{X}) \le 0\,, \tag{2.6}$$

where $A=(m-1)a/\{2(n-3)\}$. Here, let $X=(\bar{X}-\mu)/\sqrt{\sigma_1^2/m}$, $Y=(\bar{Y}-\mu)/\sqrt{\sigma_2^2/n}$, $T_1=mS_1/\sigma_1^2$ and $T_2=nS_2/\sigma_2^2$. It is easy to see that X, Y, T_1 and T_2 are mutually independent random variables such that X and Y have standard normal distributions, and T_1 and T_2 have chi-square distributions with $m-1$ and $n-1$ degrees of freedom, respectively. Let $\tau=n\sigma_1^2/(m\sigma_2^2)$ and $Z=\tau+AT_2/T_1$. Then the inequality (2.6) is rewritten by

$$(2Z-a\tau)\sqrt{\tau}X^2 + 2(a\tau-Z)XY - a\sqrt{\tau}Y^2 \ge 0\,,$$

which is equivalent to

$$(X,Y)\begin{pmatrix}(2Z-a\tau)\sqrt{\tau} & a\tau-Z\\ a\tau-Z & -a\sqrt{\tau}\end{pmatrix}\begin{pmatrix}X\\ Y\end{pmatrix} \ge 0\,. \tag{2.7}$$

There exists an orthogonal matrix P such that

$$P\begin{pmatrix}(2Z-a\tau)\sqrt{\tau} & a\tau-Z\\ a\tau-Z & -a\sqrt{\tau}\end{pmatrix}P' = \mathrm{diag}\,(\lambda_1,\lambda_2)\,,$$

where

$$\lambda_1=\{(2Z-a\tau-a)\sqrt{\tau}+[(2Z-a\tau-a)^2\tau+4Z^2]^{1/2}\}/2\,,$$
$$\lambda_2=\{(2Z-a\tau-a)\sqrt{\tau}-[(2Z-a\tau-a)^2\tau+4Z^2]^{1/2}\}/2\,.$$

Since $(X,Y)'$ has a bivariate normal distribution $N_2(\mathbf{0},\boldsymbol{I})$, so does $P(X,Y)'$, being independent of the random variable Z. Letting $(U_1,U_2)'=P(X,Y)'$, we can express (2.7) as $\lambda_1U_1^2+\lambda_2U_2^2\ge 0$, or $\lambda_1^2U_1^2+\lambda_1\lambda_2U_2^2\ge 0$, which yields

$$U_2^2/U_1^2 \le (G+\sqrt{G^2+1})^2,$$

where

$$G=\sqrt{\tau}\left\{1-\frac{a(1+\tau)}{2(\tau+AT_2/T_1)}\right\}.$$

In this way, (2.4) holds if

$$h(\tau)\overset{\text{def}}{=}P\{V\le(G+\sqrt{G^2+1})^2\}\ge 1/2,$$

uniformly, where V is a random variable having an F-distribution with degrees of freedom $(1,1)$. Letting $f_{1,1}(v)$ be a density of V, we can represent $h(\tau)$ as

$$h(\tau)=E\left[\int_0^{(G+\sqrt{G^2+1})^2} f_{1,1}(v)\,dv\right],$$

where the expectation $E[\cdot]$ is taken with respect to the random variable T_2/T_1. The dominated convergence theorem gives that $h(\tau)\to E\left[\int_0^1 f_{1,1}(v)\,dv\right]$ $=1/2$ as $\tau\to 0$, so that it is sufficient to show that $h'(\tau)$, the derivative with respect to τ, is nonnegative. Then

$$h'(\tau)=C_0E\left[\frac{(G+\sqrt{G^2+1})^{-1}}{1+(G+\sqrt{G^2+1})^2}(G+\sqrt{G^2+1})(1+G/\sqrt{G^2+1})G'\right],$$

where C_0 is a positive constant and

$$G'=\{2Z^2-a(1+3\tau)Z+2a\tau(1+\tau)\}/(4\sqrt{\tau}Z^2),$$

for $Z=\tau+AT_2/T_1$. Noting that $1+(G+\sqrt{G^2+1})^2=2(G+\sqrt{G^2+1})\cdot\sqrt{G^2+1}$, we have

$$\text{(2.8)}\qquad h'(\tau)=\frac{C_0}{2}E\left[\frac{G'}{G^2+1}\right]$$
$$=\frac{C_0}{2\sqrt{\tau}(1+\tau)}E\left[\frac{2Z^2-a(1+3\tau)Z+2\tau(1+\tau)a}{4Z^2-4a\tau Z+a^2\tau(1+\tau)}\right].$$

For $a\le 4/3$, observe that

$$\text{(2.9)}\qquad \frac{2Z^2-a(1+3\tau)Z+2\tau(1+\tau)a}{4Z^2-4a\tau Z+a^2\tau(1+\tau)}\ge\frac{1}{4}\left(2-\frac{a(1+\tau)}{Z}\right).$$

Then from (2.8) and (2.9), it follows that $h'(\tau) \geq 0$ if $E[2 - a(1+\tau)/Z] \geq 0$ or

$$a \leq 2 \inf_{\tau>0} \left\{ \frac{1}{E[(1+\tau)/Z]} \right\} \quad \text{for} \quad a \leq 4/3 . \tag{2.10}$$

Lemma 2.1 gives that

$$\inf_{\tau>0} \left\{ \frac{1}{E[(1+\tau)/Z]} \right\} \geq \min \left\{ 1, \frac{A}{E[T_1/T_2]} \right\} .$$

Since $E[T_1/T_2] = (m-1)/(n-3)$ and $A = (m-1)a/\{2(n-3)\}$, the r.h.s. of (2.10) is bounded below by $\min(2, a)$, which is greater than or equal to a for $a \leq 4/3$. Therefore the proof is complete.

Remark. Theorem 2.1 does not include a condition on c because it is proved without using information of $(\bar{X} - \bar{Y})^2$. When $c > 0$, another technique of the proof may be desirable to provide more precise conditions.

Acknowledgements

The author is grateful to Professor N. Sugiura for his valuable comments. He is also grateful to the referees for their valuable comments and helpful suggestions.

References

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

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.

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*a*). Estimation of a common mean of two normal distributions, *Tsukuba J. Math.*, **11**, 157–175.

Kubokawa, T. (1987*b*). Admissible minimax estimation of a common mean of two normal populations, *Ann. Statist.*, **15**, 1245–1256.

Peddada, S. D. and Khattree, R. (1986). On Pitman nearness and variance of estimators, *Comm. Statist. A—Theory Methods*, **15**, 3005–3017.

Pitman, E. J. G. (1937). The closest estimates of statistical parameters, *Proc. Cambridge Phil. Soc.*, **33**, 212–222.

Rao, C. R., Keating, J. P. and Mason, R. L. (1986). The Pitman nearness criterion and its determination, *Comm. Statist. A—Theory Methods*, **15**, 3173–3191.

Sen, P. K. (1986). Are BAN estimators the Pitman-closest ones too?, *Sankhyā Ser. A*, **48**, 51–58.

Sen, P. K., Kubokawa, T. and Saleh, A. K. Md. E. (1989). The Stein paradox in the sense of the Pitman measure of closeness, to appear in *Ann. Statist.*, **17**, September.

Sugiura, N. (1984). Asymptotically closer estimators for the normal covariance matrix, *J. Japan Statist. Soc.*, **14**, 145–155.