TSUKUBA J. MATH.
Vol. 11 No 1 (1987). 157—175

ESTIMATION OF A COMMON MEAN OF TWO NORMAL DISTRIBUTIONS

By

Tatsuya KUBOKAWA

Consider the problem of estimating the common mean of two normal distributions with independent estimators for variances. The paper gives sufficient conditions for the combined estimator being better than the uncombined estimator in the sense of making its variance smaller. They are extensions of some parts of the conditions by Brown and Cohen [4], Khatri and Shah [9] and Bhattacharya [1, 2]. Applications to the problem of recovery of interblock information in the BIB designs and the problem of estimating common coefficients of two regression models are shown.

1. Introduction.

The problem of estimating a common mean of two normal distributions with unknown variances has been studied in several papers. Of these, Graybill and Deal [7] showed that the necessary and sufficient condition for the combined estimator to have a smaller variance than each sample mean is the sample sizes being greater than 10. Later this is corrected by Khatri and Shah [9] as $(n_i-3)(n_j-9)\geqq 16$ for $i\neq j$, where n_1 and n_2 are sample sizes of the populations. This result has been generalized in various forms by Brown and Cohen [4], Khatri and Shah [9] and Bhattacharya [1, 2]. In this paper, assuming the underlying model by Bhattacharya [2], we extend the class of combined estimators by adding one more arbitrary constant and give sufficient conditions for the variance of the estimator being uniformly smaller than that of the uncombined estimator.

In Section 2, we give a sufficient condition based on Brown and Cohen [4] and other sufficient conditions based on the inequality of Bhattacharya [3]. Further from the inequality, we get a new sufficient condition under additional constraints on sample sizes and constant multipliers. This sufficient condition is an extended form of Bhattacharya [2] except for some special type of estimators and is proved to be better under those constraints. In Section 3, the proofs of the results in Section 2 are given.

Received April 25, 1986.

In Section 4, we specialize these results to the problem of estimating a common mean from two normal populations and apply to the problem of the recovery of interblock information in the balanced incomplete block designs. Here we give a simple sufficient condition for Yates [12]'s estimate being better than the intrablock estimate. Bhattacharya [1] obtained another sufficient condition and showed this condition was satisfied for all asymmetrical BIBD's listed in Fisher-Yates' table [6] with two exceptions. For one of these two designs, Bhattacharya [1] proved that Yates' estimate did not have the desired property, but for the other design, he could come to no conclusion. Using our sufficient condition for this design, we can see that Yates' estimate is superior to the intrablock estimate. We also apply our results to the problem of estimating common regression coefficients of two normal linear models according to Swamy and Mehta [11], where the preference of estimators is judged by usual partial ordering between covariance matrices.

2. Main results.

Let X, Y, S_1, S_2 and W_j, $j=1, \cdots, q$ be independent observed random variables where X has normal distribution $N(\mu, \alpha_0\sigma_1^2)$ and Y has $N(\mu, \beta_0\sigma_2^2)$ for known constants α_0 and β_0; S_1, S_2 and W_j are estimators for unknown parameters σ_1^2, σ_2^2 and $\alpha_j\sigma_1^2+\beta_j\sigma_2^2$ respectively with known α_j and β_j such that S_i/σ_i^2 has $\chi^2_{m_i}$-distribution ($m_i>0$), that is, chi square variate with m_i degrees of freedom for $i=1, 2$, and $W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)$ has χ_1^2-distribution for all $j=1, \cdots, q$. Let us write $q=0$ when the statistics W_j's don't exist. The problem is to find a better combined estimator than X for the unknown common mean μ within the form

$$\hat{\mu}=X+\phi\cdot(Y-X)\,, \tag{2.1}$$

where

$$\phi=\frac{a\alpha_0 S_1}{\alpha_0 S_1+c\beta_0 S_2+d\beta_0\{(Y-X)^2/\beta_0+\sum_{j=1}^{q}W_j/\beta_j\}} \tag{2.2}$$

with nonnegative constants a, c and d ($c+d>0$) suitably chosen. It is easy to see that $\hat{\mu}$ is an unbiased estimator of μ. In Section 4, the estimator $\hat{\mu}$ is applied to the problem of recovery of interblock information in BIBD's with prior knowledge that $\sigma_2^2\geqq\sigma_1^2$ between unknown variances. It is also applied to the problem of estimating a common mean of two normal populations and the problem of estimating common coefficients of two regression models with no information about σ_1^2 and σ_2^2. To deal with these applications, we suppose that

$\rho>\rho_0$ where $\rho=\beta_0\sigma_2^2/(\alpha_0\sigma_1^2)$ and ρ_0 is a nonnegative known constant. The conditions for $Var(\hat{\mu})\leqq Var(X)$ for any $\rho>\rho_0$ and any $\sigma_1^2>0$ have been given in Brown and Cohen [4] for $c=d$ when $\rho_0=0$ or $\rho_0=1$; in Khatri and Shah [9] for $c=d$ or $d=0$ when $\rho_0=0$; in Bhattacharya [2] for $c=d$, $d=0$ or $c=0$ when $\rho_0=0$ and in Bhattacharya [1] for $\alpha_0=\cdots=\alpha_q$ and $\beta_0=\cdots=\beta_q$ when $\rho_0\geqq 0$. We shall look for the sufficient conditions in terms of three constants a, c and d given in (2.2) when $\rho_0\geqq 0$.

Let W_0 be a $(\alpha_0\sigma_1^2+\beta_0\sigma_2^2)\chi_3^2$-variate independent of S_1, S_2 and W_j $(j=1, \cdots, q)$. Then the following expression of the variance of $\hat{\mu}$ according to Brown and Cohen [4] and Khatri and Shah [9] is useful.

$$Var(\hat{\mu})=Var(X)+\alpha_0\sigma_1^2E[-2\bar{\phi}+(1+\rho)\bar{\phi}^2]\,, \tag{2.3}$$

where $\bar{\phi}$ is the same as ϕ in (2.2) except that $(X-Y)^2$ is replaced by W_0. Note that the distribution of $(X-Y)^2$ is $(\alpha_0\sigma_1^2+\beta_0\sigma_2^2)\chi_1^2$ and is different from W_0. From (2.3), a necessary and sufficient condition for the estimator $\hat{\mu}$ being uniformly better than X for any $\rho\geqq\rho_0$ is given by

$$a\leqq 2\cdot\inf_{\rho>\rho_0}\left\{\frac{E[r(\rho)]}{E[\{r(\rho)\}^2]}\right\}, \tag{2.4}$$

where $r(\rho)=(1+\rho)\bar{\phi}/a$.

The following two theorems are obtained from the inequality (2.4). An extension of Brown and Cohen [4] is given by the following theorem.

THEOREM 2.1. *For $m_2+q>1$ and $c, d>0$, the variance of $\hat{\mu}$ is uniformly smaller than that of X for any $\rho>\rho_0$ if $a\leqq a_{BC}(c, d; \rho_0)$ where*

$$a_{BC}(c, d; \rho_0)=\frac{2(m_2+q+3)/(m_2+q+1)}{E\left[\max\left\{\frac{1+\rho_0}{V+\rho_0 f(c, d)}, \frac{1}{f(c, d)}\right\}V^2\right]}, \tag{2.5}$$

$f(c, d)=\{\min(c, d)\}(m_2+q+3)/m_1$ and V is a random variable having F-distribution with (m_1, m_2+q+3) degrees of freedom.

The assumption $m_2+q>1$ implies that the denominator of the r. h. s. of (2.5) is finite for any $c, d>0$. Putting $c=d=m_1/(m_2+3)$ and $q=0$ in Theorem 2.1, we get $a\leqq 2(m_2+3)/\{(m_2+1)E[\max(V, V^2)]\}$ for $\rho_0=0$ and $a\leqq 2(m_2+3)/\{(m_2+1)E[\max(2/(V+1), 1)V^2]\}$ for $\rho_0=1$, because $f(c, d)=1$. These were derived by Brown and Cohen [4].

Next we define two random variables Z and T such that

$$Z=\frac{\sum_{j=1}^{q}\chi_j^2(1)+\chi_0^2(3)}{\chi^2(m_2)+\sum_{j=1}^{q}\chi_j^2(1)+\chi_0^2(3)}, \tag{2.6}$$

(2.7)
$$T=\frac{\sum_{j=1}^{q}\{\beta_0\alpha_j/(\alpha_0\beta_j)\}\chi_j^2(1)+\chi_0^2(3)}{\chi^2(m_2)+\sum_{j=1}^{q}\chi_j^2(1)+\chi_0^2(3)},$$

where $\chi^2(m_2)$, $\chi_0^2(3)$, $\chi_1^2(1), \cdots, \chi_q^2(1)$ are mutually independent χ^2-variates with degrees of freedom shown in the parentheses respectively. We note that Z follows beta distribution with parameters $((q+3)/2, m_2/2)$ and that Z and T are not independent. An extension of Bhattacharya [1, 2] is given by the following theorem.

THEOREM 2.2. *Suppose that one of the following three conditions holds:* (i) $m_2+q>1$ *if* $c, d>0$, (ii) $m_2>4$ *if* $d=0$ *or* (iii) $q>1$ *if* $c=0$. *Put* $a_0=(m_2+q-1)/(m_1+2)$. *Then the variance of* $\hat{\mu}$ *is uniformly smaller than that of* X *for any* $\rho>\rho_0$ *if* $a\leqq 2\max\left[\min\left\{1, \inf_{\rho>\rho_0} A(c, d; 1/\rho)a_0\right\}, \inf_{\rho>\rho_0} A(c, d; 1/\rho)a_0\rho_0/(1+\rho_0)\right]$ *where*

(2.8)
$$A(c, d; 1/\rho)=\frac{E[\{c(1-Z)+dZ+dT/\rho\}^{-1}]}{E[\{c(1-Z)+dZ+dT/\rho\}^{-2}]}.$$

The three assumptions on m_2 and q in Theorem 2.2 mean that $0<A(c, d; 1/\rho)<\infty$ and $a_0>0$. The sufficient condition in Theorem 2.2 is not simple to be checked since it contains the infimum and the expectations. We shall give weaker versions of it, which are, however, more useful. Since

(2.9)
$$E[f(X)]/E[g(X)]\geqq\inf_x\{f(x)/g(x)\}$$

holds for any positive valued functions f, g and any random variable X, it follows that

(2.10)
$$A(c, d; 1/\rho)\geqq\inf_{t>0,\, 0<z<1}\{c(1-z)+dz+dt/\rho\}$$
$$=\min(c, d),$$

which yields

COROLLARY 2.1. *For* $m_2+q>1$ *and* $c, d>0$, *the estimator* $\hat{\mu}$ *is uniformly better than* X *for any* $\rho>\rho_0$ *if*

$$a\leqq 2\max[\min\{1, \min(c, d)a_0\}, \min(c, d)a_0\rho_0/(1+\rho_0)].$$

The sufficient condition shown in Corollary 2.1 was proved by Bhattacharya [1] for $\alpha_0=\cdots=\alpha_q$ and $\beta_0=\cdots=\beta_q$ when $\rho_0=0$.

Furthermore from Theorem 2.2, we can develop more precise sufficient conditions, which are used in all the applications in Section 4. For this, we assume the following conditions:

(C-1) $$c \geqq d \geqq 0 \quad \text{or} \quad c=0 \quad (c+d>0),$$

(C-2) $$2 \min_{0 \leq j \leq q} \left\{ \frac{\alpha_j}{\beta_j} \right\} \geqq \max_{0 \leq j \leq q} \left\{ \frac{\alpha_j}{\beta_j} \right\}.$$

For $c=d$ or $d=0$, however, we need not assume the condition (C-2). We note that the condition (C-1) includes three cases $c=d$, $d=0$ or $c=0$ which have been studied by Brown and Cohen [4], Khatri and Shah [9] and Bhattacharya [2], and that the condition (C-2) is always satisfied for $q=0$ or for $\alpha_0=\cdots=\alpha_q$, $\beta_0=\cdots=\beta_q$. Under these conditions (C-1) and (C-2), we can show that $\inf_{\rho>\rho_0} A(c, d; 1/\rho)=A(c, d; 0)$ in Theorem 2.2, and get the next theorem.

THEOREM 2.3. *Suppose that the condition* (C-1) *holds, and that the condition* (C-2) *holds except when $c=d$ or $d=0$. Assume that one of the following three conditions is satisfied:* (i) *$m_2+q>1$ if c, $d>0$,* (ii) *$m_2>4$ if $d=0$ or* (iii) *$q>1$ if $c=0$. Put $a_0=(m_2+q-1)/(m_1+2)$ and*

(2.11) $$A(c, d; 0)=\frac{E[\{c(1-Z)+dZ\}^{-1}]}{E[\{c(1-Z)+dZ\}^{-2}]},$$

where Z has beta distribution with parameters $((q+3)/2, m_2/2)$. Then

(a) *$\hat{\mu}$ is uniformly better than X for any $\rho>\rho_0$ if*

$$a \leqq 2 \max[\min\{1, A(c, d; 0)a_0\}, A(c, d; 0)a_0\rho_0/(1+\rho_0)].$$

(b) *Given $A(c, d; 0)a_0 \leqq 1$, $\hat{\mu}$ is uniformly better than X for any $\rho>\rho_0$ if and only if $a \leqq 2A(c, d; 0)a_0$.*

(c) *Given $a \leqq 2$, $\hat{\mu}$ is uniformly better than X for any $\rho>\rho_0$ if and only if $A(c, d; 0) \geqq a/(2a_0)$.*

(d) *$\hat{\mu}$ is uniformly better than X for any $\rho>\rho_0$ if*

(2.12) $$a \leqq 2 \max\left[\min\left\{1, \left(1+\frac{m_2(c-d)}{(m_2+q+3)c}\right)da_0\right\}, \left(1+\frac{m_2(c-d)}{(m_2+q+3)c}\right)\frac{da_0\rho_0}{1+\rho_0}\right] \quad \textit{for } c \geqq d>0,$$

or if

(2.13) $$a \leqq 2 \max\left[\min\left\{1, \frac{m_2-4}{m_1+2}c\right\}, \frac{(m_2-4)c\rho_0}{(m_1+2)(1+\rho_0)}\right] \quad \textit{for } c>0,\ m_2>4.$$

The assumptions on m_2 and q in Theorem 2.3 guarantee the existence of the expectations in $A(c, d; 0)$ in (2.11) and $a_0>0$, which are equal to those of Theorem 2.2. Special cases of Theorem 2.3 when $A(c, c; 0)=c$, $A(c, 0; 0)=c(m_2-4)/(m_2+q-1)$ or $A(0, d; 0)=d(q-1)/(m_2+q-1)$ with $\rho_0=0$ were proved by Bhattacharya [2] without assuming the condition (C-2). However, we should impose (C-2) for $c=0$ in our proof. In order to compute the upper bound of a

in Theorem 2.3, we can rewrite $A(c, d; 0)$ according to Khatri and Shah [9] as

$$A(c, d; 0)=\frac{{}_2F_1\left(1, \frac{q+3}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right)}{{}_2F_1\left(2, \frac{q+3}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right)}c \tag{2.14}$$

where $c\geqq d\geqq 0$ and ${}_2F_1$ is the hypergeometric function. When $d=0$, the assumption $m_2>4$ implies that the hypergeometric functions in (2.14) converge. The simple sufficient condition (2.12) is derived from the result (a) and (2.14), and is useful for the constant d away from zero. The other sufficient condition (2.13) is independent of d and the r. h. s. is equal to the upper bound of a in Theorem 2.3 (a) when $d=0$ and is smaller when $d=c$.

If $\rho_0=0$ and the conditions (C-1), (C-2) hold, the sufficient condition (a) given in Theorem 2.3 is better than that in Theorem 2.1 as is shown in

THEOREM 2.4. *Suppose that one of the following conditions holds:* (i) $m_2+q>1$ *if* $c, d>0$, (ii) $m_2>4$ *if* $d=0$ *or* (iii) $q>1$ *if* $c=0$. *If* $\rho_0=0$, *then we get the following inequality between two upper bounds of* a *given in Theorems* 2.1 *and* 2.3 (*a*).

$$a_{BC}(c, d; 0)\leqq 2\min\{1, A(c, d; 0)a_0\} \tag{2.15}$$

for any nonnegative constants c *and* d *not all equal to zero. The inequality holds without assuming the conditions* (C-1) *and* (C-2).

3. Proofs of theorems.

To prove main theorems in Section 2, we shall express the random variable $r(\rho)$ in (2.4) by other random variables whose distributions are independent of unknown parameters. Using the observations S_1, S_2, W_j $(j=1, \cdots, q)$ and the random variable W_0 defined in Section 2, put

$$F=\frac{S_2/\sigma_2^2+\sum_{j=0}^q W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)}{S_1/\sigma_1^2}, \tag{3.1}$$

$$Z=\frac{\sum_{j=0}^q W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)}{S_2/\sigma_2^2+\sum_{j=0}^q W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)}, \tag{3.2}$$

$$T=\frac{\sum_{j=0}^q\{\beta_0\alpha_j/(\alpha_0\beta_j)\}W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)}{S_2/\sigma_2^2+\sum_{j=0}^q W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)}. \tag{3.3}$$

It is easy to see that $\{m_1/(m_2+q+3)\}F$ has F-distribution with (m_2+q+3, m_1) degrees of freedom and Z has beta distribution with parameters $((q+3)/2, m_2/2)$. It follows that the distribution of Z and T are given by (2.6) and (2.7). Since

$\sum_{i=1}^{m}\chi^2_{a_i}$ and $(\chi^2_{a_1}/\sum_{i=1}^{m}\chi^2_{a_i}, \cdots, \chi^2_{a_m}/\sum_{i=1}^{m}\chi^2_{a_i})$ are independent for the independent χ^2-variates with a_i degrees of freedoms and a natural number m, we can see that F and (Z, T) are independent.

Now we express $r(\rho)$ in (2.4) by the random variables F, Z and T. We first see that

$$r(\rho)=(1+\rho)\bar{\phi}/a = \frac{(1+\rho)\alpha_0 S_1}{\alpha_0 S_1 + c\beta_0 S_2 + d\beta_0 \sum_{j=0}^{q} W_j/\beta_j}. \tag{3.4}$$

Note that

$$\begin{aligned} &\textstyle\sum_{j=0}^{q} W_j/(\alpha_j\sigma_1^2+\beta_j\sigma_2^2) = ZF\,S_1/\sigma_1^2, \\ &S_2/\sigma_2^2 = (1-Z)F\,S_1/\sigma_1^2, \\ &\textstyle\sum_{j=0}^{q} \frac{\beta_0\alpha_j W_j}{\alpha_0\beta_j(\alpha_j\sigma_1^2+\beta_j\sigma_2^2)} = TF\,S_1/\sigma_1^2. \end{aligned} \tag{3.5}$$

Since $\rho=\beta_0\sigma_2^2/(\alpha_0\sigma_1^2)$, the denominator in the last expression of (3.4) can be rewritten as

$$\begin{aligned} &\alpha_0 S_1 + c\beta_0 S_2 + \alpha_0\sigma_1^2 d \textstyle\sum_{j=0}^{q}\left(\frac{\beta_0\alpha_j}{\alpha_0\beta_j}+\rho\right)\frac{W_j}{\alpha_j\sigma_1^2+\beta_j\sigma_2^2} \\ &= \alpha_0 S_1\{1+\rho c(1-Z)F + d(TF+\rho ZF)\} \\ &= \alpha_0 S_1[1+dTF+\rho\{c(1-Z)+dZ\}F]. \end{aligned} \tag{3.6}$$

Hence $r(\rho)$ is represented in the form

$$r(\rho)=(1+\rho)/[1+dTF+\rho\{c(1-Z)+dZ\}F]. \tag{3.7}$$

Putting $\rho=\rho_0$ or $\rho=\infty$ in (3.7) and using (2.4) with

$$\begin{aligned} &E[F^{-1}]=m_1/(m_2+q+1), \\ &E[F^{-2}]=m_1(m_1+2)/\{(m_2+q-1)(m_2+q+1)\}, \end{aligned} \tag{3.8}$$

we get the following necessary condition, which is used to prove Theorems 2.3 and 2.4.

LEMMA 3.1. *Suppose that one of the following conditions holds:* (i) $m_2+q>1$ *if* $c, d>0$, (ii) $m_2>4$ *if* $d=0$ *or* (iii) $q>1$ *if* $c=0$. *Then a necessary condition for* $\hat{\mu}$, *given by* (2.1), *being uniformly better than* X *for any* $\rho>\rho_0$ *is*

$$a \leqq 2 \min\left\{\frac{E[r(\rho_0)]}{E[\{r(\rho_0)\}^2]},\ A(c, d; 0)a_0\right\} \tag{3.9}$$

where $A(c, d; 0)$ *and* a_0 *are given by Theorem* 2.3.

The three assumptions on m_2 and q in Lemma 3.1 guarantee that all the expectations in (2.4) and (3.9) exist. In fact, it is enough to show that $E[\{c(1-Z)+dZ\}^{-2}]E[F^{-2}]<\infty$. It is easily seen that

$$(3.10)\qquad E[\{c(1-Z)+dZ\}^{-2}]\begin{cases} \leqq \max(c^{-2}, d^{-2}) & \text{for } c,\ d>0, \\ =E[(1-Z)^{-2}]/c^2 & \text{for } d=0, \\ =E[Z^{-2}]/d^2 & \text{for } c=0, \end{cases}$$

and that $E[F^{-2}]<\infty$ for $m_2+q>1$ from (3.8). Noting that $E[(1-Z)^{-2}]<\infty$ for $m_2>4$ and $E[Z^{-2}]<\infty$ for $q>1$, we have the three assumptions on m_2 and q in Lemma 3.1.

3.1. Proof of Theorem 2.1. We note that all the expectations in the proof exist for $c,\ d>0$ if $m_2+q>1$ as shown by the similar discussion in Lemma 3.1. Following Brown and Cohen [4], consider $r(\rho)$ in (3.7). Then $\sup_{\rho>\rho_0}\{r(\rho)\}\leqq\max\{r(\rho_0),\ r(\infty)\}$, which yields

$$(3.11)\qquad \begin{aligned} r(\rho) &\leqq \max\left\{\frac{1+\rho_0}{1+\rho_0\{c(1-Z)+dZ\}F},\ \frac{1}{\{c(1-Z)+dZ\}F}\right\} \\ &\leqq \max\left\{\frac{1+\rho_0}{1+\rho_0\min(c,\ d)F},\ \frac{1}{\min(c,\ d)F}\right\} \\ &= h(F) \qquad \text{(say)}. \end{aligned}$$

Then we have

$$(3.12)\qquad \frac{E[r(\rho)]}{E[\{r(\rho)\}^2]}\geqq\frac{E[r(\rho)]}{E[h(F)r(\rho)]}\geqq\inf_{\substack{\rho>\rho_0,\ t>0\\ 0<z<1}}\left\{\frac{E[r(\rho)\mid T=t,\ Z=z]}{E[h(F)r(\rho)\mid T=t,\ Z=z]}\right\}$$

by (2.9). When $c=d$, Brown and Cohen [4] showed that the bracketed term on the r. h. s. of (3.12) is nonincreasing in ρ and that the infimum is attained when $\rho\to\infty$. This fact can be similarly shown to be true without assuming $c=d$. However, in this place, we directly prove based on Bhattacharya [3] that

$$(3.13)\qquad \frac{E[r(\rho)\mid T=t,\ Z=z]}{E[h(F)r(\rho)\mid T=t,\ Z=z]}\geqq\frac{E[F^{-1}]}{E[h(F)F^{-1}]},$$

for any $\rho>\rho_0$, $t>0$ and $0<z<1$. The r. h. s. of (3.13) is obtained by letting $\rho\to\infty$ in the l. h. s.. If the inequality (3.13) is valid, then putting $F^{-1}=\{m_1/(m_2+q+3)\}V$ in the denominator of the r. h. s. of (3.13) and noting (3.8), (3.12) and (2.4) gives the sufficient condition $a\leqq a_{BC}(c,\ d\,;\rho_0)$ in Theorem 2.1. So we shall prove (3.13). The independence between F and $(T,\ Z)$ first implies that the inequality (3.13) can be rewritten in the form

(3.14) $\qquad E_0[Fr(\rho) \mid T=t,\ Z=z]\ E_0[h(F)] \geqq E_0[h(F)Fr(\rho) \mid T=t,\ Z=z]$

where E_0 stands for expectation with respect to the probability measure P_0 given by $P_0(A)=E[I_A F^{-1}]/E[F^{-1}]$ and I_A is the indicator function of a set A. Regarding $r(\rho)$ in (3.7) as a function of F, it is easy to see that $Fr(\rho)$ is nondecreasing in F given T, Z and that $h(F)$ is decreasing in F. Hence we get the inequality (3.14), which completes the proof.

3.2. Proof of Theorem 2.2. Note that the random variable $r(\rho)$ in (3.7) is represented as $r(\rho)=\{(1-\Theta)+\Theta R\}^{-1}$ where $\Theta=\rho/(1+\rho)$ and $R=\{c(1-Z)+dZ+dT/\rho\}F$. Then it follows from the inequality in Bhattacharya [3, theorem 2.2] that

(3.15) $$\frac{E[r(\rho)]}{E[\{r(\rho)\}^2]} \geqq \min\left\{1,\ \frac{E[R^{-1}]}{E[R^{-2}]}\right\}.$$

On the other hand when $\rho_0>0$, we have

(3.16) $$\frac{E[r(\rho)]}{E[\{r(\rho)\}^2]} = \Theta\frac{E[\{1/\rho+R\}^{-1}]}{E[\{1/\rho+R\}^{-2}]} \geqq \frac{\rho_0}{1+\rho_0}\cdot\frac{E[R^{-1}]}{E[R^{-2}]},$$

because $\Theta>\rho_0/(1+\rho_0)$ and the following inequality holds:

(3.17) $$\frac{E[\{1/\rho+R\}^{-1}]}{E[\{1/\rho+R\}^{-2}]} \geqq \frac{E[R^{-1}]}{E[R^{-2}]}.$$

In fact, this is equivalent to the inequality

(3.18) $$E_1\left[\frac{R}{1/\rho+R}\right]E_1\left[\frac{1}{R}\right] \geqq E_1\left[\frac{R}{\{1/\rho+R\}^2}\right],$$

where $E_1[\cdot]$ is the expectation according to the probability measure P_1 given by $P_1(A)=E[I_A R^{-1}]/E[R^{-1}]$ and I_A is the indicator function of a set A. Since $R/(1/\rho+R) \leqq 1$, it is enough to show that

(3.19) $$E_1\left[\frac{R}{1/\rho+R}\right]E_1\left[\frac{1}{R}\right] \geqq E_1\left[\frac{1}{1/\rho+R}\right],$$

which is proved because $R/(1/\rho+R)$ is increasing in R and $1/R$ is decreasing in R. Hence we get the inequality (3.16). Here, since F is independent of (T, Z), we can see that

(3.20) $$\frac{E[R^{-1}]}{E[R^{-2}]} = \frac{E[\{c(1-Z)+dZ+dT/\rho\}^{-1}]}{E[\{c(1-Z)+dZ+dT/\rho\}^{-2}]}\cdot\frac{E[F^{-1}]}{E[F^{-2}]} = A(c,\ d\,;1/\rho)a_0,$$

from (3.8) where $A(c, d; 1/\rho)$ and a_0 are defined in Theorem 2.2. Combining two inequalities (3.15), (3.16) and noting (2.4), we get the sufficient condition $a \leqq 2\max[\min\{1, \inf_{\rho>\rho_0} A(c, d; 1/\rho)a_0\}, \inf_{\rho>\rho_0} A(c, d; 1/\rho)a_0\rho_0/(1+\rho_0)]$ in Theorem 2.2. We imposed the assumptions on m_2 and q in Theorem 2.2 by the similar discussion in Lemma 3.1 in order that the expectations in this proof are finite.

3.3. Proof of Theorem 2.3. To prove (a), we show that

$$\inf_{\rho>\rho_0} A(c, d; 1/\rho) = A(c, d; 0) \tag{3.21}$$

in Theorem 2.2 under the conditions (C-1) and (C-2). We see from Theorem 2.2 that $A(c, d; 1/\rho)$ in (2.8) and a_0 are positive for any $\rho > \rho_0$ if the assumptions on m_2 and q in Theorem 2.3 are satisfied. From (2.8) and (2.11), the equation (3.21) is equivalent to

$$\begin{aligned} &E_2\left[\frac{c(1-Z)+dZ}{c(1-Z)+dZ+dT/\rho}\right]E_2\left[\frac{1}{c(1-Z)+dZ}\right] \\ &\geqq E_2\left[\frac{c(1-Z)+dZ}{\{c(1-Z)+dZ+dT/\rho\}^2}\right], \end{aligned} \tag{3.22}$$

for any $\rho > \rho_0$, where $E_2[\cdot]$ stands for expectation with respect to the probability measure P_2 given by $P_2(A) = E[I_A\{c(1-Z)+dZ\}^{-1}]/E[\{c(1-Z)+dZ\}^{-1}]$ and I_A is the indicator function of a set A. The inequality (3.22) is evident for $c=d$ or $d=0$, so that from the condition (C-1), we prove (3.22) in the case of $c>d>0$ or $c=0$. Put $\underline{\gamma} = \min_{0\leq j\leq q}\{\beta_0\alpha_j/(\alpha_0\beta_j)\}$ and $\bar{\gamma} = \max_{0\leq j\leq q}\{\beta_0\alpha_j/(\alpha_0\beta_j)\}$. Then we get $\underline{\gamma}Z \leqq T \leqq \bar{\gamma}Z$ in (3.2) and (3.3), so that it suffices to show that

$$\begin{aligned} &E_2\left[\frac{c(1-Z)+dZ}{c(1-Z)+dZ+d\bar{\gamma}Z/\rho}\right]E_2\left[\frac{1}{c(1-Z)+dZ}\right] \\ &\geqq E_2\left[\frac{c(1-Z)+dZ}{\{c(1-Z)+dZ+d\underline{\gamma}Z/\rho\}^2}\right]. \end{aligned} \tag{3.23}$$

Note that the integrand in the r. h. s. of (3.23) is bounded from above by $\{c(1-Z)+dZ+d\bar{\gamma}Z/\rho\}^{-1}$, because of $2\underline{\gamma} \geqq \bar{\gamma}$ by the condition (C-2). Hence it is enough to show that

$$\begin{aligned} &E_2\left[\frac{c(1-Z)+dZ}{c(1-Z)+dZ+d\bar{\gamma}Z/\rho}\right]E_2\left[\frac{1}{c(1-Z)+dZ}\right] \\ &\geqq E_2\left[\frac{1}{c(1-Z)+dZ+d\bar{\gamma}Z/\rho}\right]. \end{aligned} \tag{3.24}$$

For $c=0$, the equality holds in (3.24). For $c>d>0$, $1/\{c(1-Z)+dZ\}$ is increasing in Z and $\{c(1-Z)+dZ\}/\{c(1-Z)+dZ+d\bar{\gamma}Z/\rho\}$ is decreasing in Z, so that the inequality (3.24) holds. This proves (a).

When $A(c, d; 0)a_0 \leqq 1$ or $a \leqq 2$, the sufficient condition (a) in Theorem 2.3 becomes $a \leqq 2A(c, d; 0)a_0$, which is equivalent to the necessary condition in Lemma 3.1 because the following inequality holds:

$$(3.25) \qquad A(c, d; 0)a_0 \leqq \frac{E[r(\rho_0)]}{E[\{r(\rho_0)\}^2]} \quad \textit{for } \rho_0 > 0 .$$

As a matter of fact, this follows since $a \leqq 2A(c, d; 0)a_0$ is sufficient and $a \leqq 2E[r(\rho_0)]/E[\{r(\rho_0)\}^2]$ is necessary, which completes the proof of Theorem 2.3 (b) and (c).

We shall prove (d) from (a). At first using the expression (2.14), we shall derive a sufficient condition (2.12). Note that for any real values α, β, γ and $0<x<1$,

$$(3.26) \qquad {}_2F_1(\alpha, \beta; \gamma; x) = (1-x)^{\gamma-\alpha-\beta}{}_2F_1(\gamma-\alpha, \gamma-\beta; \gamma; x),$$

$$(3.27) \qquad {}_2F_1(\alpha+1, \beta; \gamma; x) = {}_2F_1(\alpha, \beta; \gamma; x) + (\beta/\gamma)x \cdot {}_2F_1(\alpha+1, \beta+1; \gamma+1; x).$$

The first equation is from Exton [5] and the second equation is obtained just by rearrangement of the coefficients in the infinite series in the l. h. s.. Then (2.14) is written by

$$(3.28) \qquad A(c, d; 0) = \frac{\left(\frac{d}{c}\right)^{m_2/2-1} {}_2F_1\left(\frac{m_2+q-1}{2}+1, \frac{m_2}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right)}{\left(\frac{d}{c}\right)^{m_2/2-2} {}_2F_1\left(\frac{m_2+q-1}{2}, \frac{m_2}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right)} c$$

$$= d\left\{1 + \frac{m_2(c-d)\cdot {}_2F_1\left(\frac{m_2+q-1}{2}+1, \frac{m_2}{2}+1; \frac{m_2+q+3}{2}+1; \frac{c-d}{c}\right)}{(m_2+q+3)c\cdot {}_2F_1\left(\frac{m_2+q-1}{2}, \frac{m_2}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right)}\right\}.$$

Evaluation of each term in the infinite series gives

$$(3.29) \qquad {}_2F_1\left(\frac{m_2+q-1}{2}+1, \frac{m_2}{2}+1; \frac{m_2+q+3}{2}+1; \frac{c-d}{c}\right) \geqq {}_2F_1\left(\frac{m_2+q-1}{2}, \frac{m_2}{2}; \frac{m_2+q+3}{2}; \frac{c-d}{c}\right),$$

which yields $A(c, d; 0) \geqq d[1+m_2(c-d)/\{(m_2+q+3)c\}]$. Hence we get the sufficient condition (2.12). Next using (2.11) and the inequality of Bhattacharya [3], we have for $c > d \geqq 0$,

(3.30)
$$A(c, d; 0)=c\frac{E[\{d/c+(1-d/c)(1-Z)\}^{-1}]}{E[\{d/c+(1-d/c)(1-Z)\}^{-2}]}$$
$$\geqq c \min\left\{1, \frac{E[(1-Z)^{-1}]}{E[(1-Z)^{-2}]}\right\}$$
$$=c\frac{m_2-4}{m_2+q-1}.$$

From (a), we obtain the sufficient condition (2.13).

3.4. Proof of Theorem 2.4. Since $\rho_0=0$, it is enough to show that

(3.31)
$$a_{BC}(c, d; 0)\leqq 2 \min\{1, A(c, d; 0)a_0\},$$

for nonnegative constants c and d $(c+d>0)$. We note from Theorem 2.2 that $A(c, d; 0)$ in (2.11) and a_0 are positive if the assumptions on m_2 and q in Theorem 2.4 are satisfied. When $c=0$ or $d=0$, then $a_{BC}(c, d; 0)=0$ and the inequality (3.31) holds. When $c, d>0$, we shall check the following two cases.

Case 1. $A(c, d; 0)a_0\leqq 1$. Given any a such that $a\leqq a_{BC}(c, d; 0)$, $\hat{\mu}$ has a smaller variance than X by Theorem 2.1. Then a should satisfy $a\leqq 2A(c, d; 0)a_0$ by Lemma 3.1. Hence we get the inequality $a_{BC}(c, d; 0)$ $\leqq 2A(c, d; 0)a_0$, which is less than 2, establishing (3.31).

Case 2. $A(c, d; 0)a_0>1$. In this case, the r. h. s. of (3.31) is equal to 2. We also see that $a_{BC}(c, d; 0)\leqq 2$, because $E[\max\{V, \max(1/c, 1/d)m_1V^2/(m_2+q+3)\}]\geqq E[V]=(m_2+q+3)/(m_2+q+1)$. Therefore the inequality (3.31) holds. Thus in all cases the proof is complete.

4. Applications.

4.1. Estimation of a common mean. Let $(X_1, \cdots, X_m)$ and $(Y_1, \cdots, Y_n)$ be independent random samples from two normal populations having a common unknown mean μ and unknown variances σ_x^2 and σ_y^2 respectively. Let $\bar{X}=\sum_{i=1}^m X_i/m$, $S_x=\sum_{i=1}^m(X_i-\bar{X})^2$ and $\bar{Y}$, S_y be defined similarly. Let us make the following match ups $(\sim)$ with the terms used in the first paragraph of Section 2: $X\sim\bar{X}$, $Y\sim\bar{Y}$, $S_1\sim S_x$, $S_2\sim S_y$, $\sigma_1^2\sim\sigma_x^2$, $\sigma_2^2\sim\sigma_y^2$, $\alpha_0\sim 1/m$, $\beta_0\sim 1/n$, $m_1\sim m-1$, $m_2\sim n-1$, $q\sim 0$. The combined estimator induced from (2.1) and (2.2) by these correspondences is

(4.1)
$$\hat{\mu}(a, c, d)=\bar{X}+\frac{aS_x/m}{S_x/m+cS_y/n+d(\bar{X}-\bar{Y})^2}(\bar{Y}-\bar{X}).$$

This includes as particular cases the estimators $T_1(a^*, c^*)$ and $T_2(a^*, c^*)$ of Bhattacharya [2]. In fact, $T_1(a^*, c^*)=\hat{\mu}(a^*, c^*(m-1)/(n-1), c^*(m-1)/(n-1))$

and $T_2(a^*, c^*)=\hat{\mu}(a^*, c^*(m-1)/(n-1), 0)$. Note that the condition (C-2) in Theorem 2.3 is always satisfied and that ρ_0 defined in Section 2 is equal to zero in this model. Using Theorem 2.3, we get the following results for $c\geqq d\geqq 0$ $(c+d>0)$: Suppose that $n>2$ for $d>0$ or $n>5$ for $d=0$. Put $A_1(c, d; 0)=E[\{c(1-Z_1)+dZ_1\}^{-1}]/E[\{c(1-Z_1)+dZ_1\}^{-2}]$ corresponding to (2.11), where Z_1 has beta distribution with parameters $(3/2, (n-1)/2)$. Then $\hat{\mu}(a, c, d)$ in (4.1) is better than $\bar{X}$ if

$$a\leqq 2\min\{1, A_1(c, d; 0)(n-2)/(m+1)\}. \tag{4.2}$$

When $A_1(c, d; 0)(n-2)/(m+1)\leqq 1$ or $a\leqq 2$, $\hat{\mu}(a, c, d)$ is better than $\bar{X}$ if and only if

$$a\leqq 2A_1(c, d; 0)(n-2)/(m+1). \tag{4.3}$$

We also obtain simple sufficient conditions $a\leqq 2\min\{1, [1+(n-1)(c-d)/\{(n+2)c\}]d(n-2)/(m+1)\}$ for $d>0$ or $a\leqq 2\min\{1, (n-5)c/(m+1)\}$ for $c>0$, $n>5$.

In particular for the estimators $T_1(a^*, c^*)$ and $T_2(a^*, c^*)$ of Bhattacharya [2], we get sufficient conditions from (4.2) as

$$a^*\leqq 2\min\left\{1, \frac{(m-1)(n-2)}{(n-1)(m+1)}c^*\right\} \quad \text{for } T_1(a^*, c^*), \tag{4.4}$$

$$a^*\leqq 2\min\left\{1, \frac{(m-1)(n-5)}{(n-1)(m+1)}c^*\right\} \quad \text{for } T_2(a^*, c^*), \tag{4.5}$$

because $A_1(c, c; 0)=c$ and $A_1(c, 0; 0)=c(n-5)/(n-2)$ for any $c>0$. These special cases were obtained by Bhattacharya [2]. He also obtained necessary and sufficient conditions derived from (4.3) for $T_1(a^*, c^*)$ and $T_2(a^*, c^*)$. From Lemma 3.1 with $d=0$, we note that the sufficient condition (4.5) is also necessary, which is better than Bhattacharya [2].

4.2. Recovery of interblock information. Consider a balanced incomplete block design (BIBD) with both blocks and errors random whose canonical form is given by Graybill and Weeks [8] as follows: Let t=number of treatments, b=number of blocks, r=number of replications per treatment, k=number of cells per block, λ=number of times any pair of treatments appears in the same block, σ^2=error variance, σ_β^2=block variance and put $f=bk-b-t+1$. The $(t-1)\times 1$ vector $\boldsymbol{x}=(x_i)$ is distributed normally with mean $\boldsymbol{\tau}=(\tau_i)$ and covariance matrix $\{k/(\lambda t)\}\sigma^2\boldsymbol{I}$ (referred to as the *intrablock estimate*), where τ_i stands for a treatment contrast. The $(t-1)\times 1$ vector $\boldsymbol{y}=(y_i)$ is distributed normally with mean $\boldsymbol{\tau}=(\tau_i)$ and covariance matrix $\{k/(r-\lambda)\}(\sigma^2+k\sigma_\beta^2)\boldsymbol{I}$ (referred to as the *interblock estimate*). The scalar S^2/σ^2 has χ_f^2-distribution and the scalar $S^{*2}/(\sigma^2+k\sigma_\beta^2)$ has χ_{b-t}^2-distribution. The total sample mean z is normally distributed with total mean ν and variance $(\sigma^2+k\sigma_\beta^2)/(bk)$. The statistics $x_1, \cdots, x_{t-1}, y_1, \cdots, y_{t-1}, S^2, S^{*2}$ and z are mutually independent. We shall assume $b>t>2$ (i. e. *asymmetrical* BIBD's) throughout this paper.

Consider the problem of estimating the common mean τ_1, which is, without loss of generality, *any treatment contrast*. Let us make the following match ups ($\sim$) with the terms used in Section 2: $X\sim x_1$, $Y\sim y_1$, $S_1\sim S^2$, $S_2\sim S^{*2}$, $\mu\sim\tau_1$, $\sigma_1^2\sim\sigma^2$, $\sigma_2^2\sim\sigma^2+k\sigma_\beta^2$, $m_1\sim f$, $m_2\sim b-t$, $q\sim t-2$, $W_j\sim(x_{j+1}-y_{j+1})^2$ ($j=1, \cdots, t-2$), $\alpha_j\sim k/(\lambda t)$ and $\beta_j\sim k/(r-\lambda)$ ($j=0, \cdots, t-2$). The combined estimate induced from (2.1) and (2.2) by these correspondences is

$$\hat{\tau}(a, c, d)=x_1+\frac{a\dfrac{k}{\lambda t}S^2}{\dfrac{k}{\lambda t}S^2+c\dfrac{k}{r-\lambda}S^{*2}+d(\boldsymbol{x}-\boldsymbol{y})'(\boldsymbol{x}-\boldsymbol{y})}(y_1-x_1). \tag{4.6}$$

This includes the estimators $T_3(a^*, c^*)$ and $T_5(a^*, c^*)$ of Bhattacharya [2] in the case of BIBD's. In fact, noting that the eigen value ϕ_i in Bhattacharya [2] corresponds to $r-\lambda$ in this model for each i, we have $T_3(a^*, c^*)=\hat{\tau}(a^*, c^*(r-\lambda)/(\lambda t), c^*(r-\lambda)/(\lambda t))$ and $T_5(a^*, c^*)=\hat{\tau}(a^*, c^*(r-\lambda)/(\lambda t), 0)$. Note that the condition (C-2) in Theorem 2.3 is satisfied for BIBD's and that $\rho_0/(1+\rho_0)$ in Theorem 2.3 is equal to $\lambda t/(rk)$ since $\rho_0=\lambda t/(r-\lambda)$ and the relation $r(k-1)=\lambda(t-1)$ in BIBD's. Then we can apply Theorem 2.3 and get

THEOREM 4.1. *Suppose that* $c\geqq d\geqq 0$ *or* $c=0$ *with* $c+d>0$ *and that one of the following three conditions holds:* (i) $b>3$ *for* $c\geqq d>0$, (ii) $b>t+4$ *for* $d=0$ *or* (iii) $t>3$ *for* $c=0$. *Put* $A_2(c, d; 0)=E[\{c(1-Z_2)+dZ_2\}^{-1}]/E[\{c(1-Z_2)+dZ_2\}^{-2}]$ *where* Z_2 *has beta distribution with parameters* $((t+1)/2, (b-t)/2)$.

(a) $\hat{\tau}(a, c, d)$ *is better than* x_1 *if* $a\leqq 2\max[\min\{1, A_2(c, d; 0)(b-3)/(f+2)\}, A_2(c, d; 0)(b-3)\lambda t/\{(f+2)rk\}]$.

(b) *Given* $A_2(c, d; 0)(b-3)/(f+2)\leqq 1$, $\hat{\tau}(a, c, d)$ *is better than* x_1 *if and only if* $a\leqq 2A_2(c, d; 0)(b-3)/(f+2)$.

(c) *Given* $a\leqq 2$, $\hat{\tau}(a, c, d)$ *is better than* x_1 *if and only if* $A_2(c, d; 0)\geqq a(f+2)/\{2(b-3)\}$.

(d) $\hat{\tau}(a, c, d)$ *is better than* x_1 *if*

$$a\leqq 2\max\Big[\min\Big\{1, \Big(1+\frac{(b-t)(c-d)}{(b+1)c}\Big)\frac{d(b-3)}{f+2}\Big\}, \Big(1+\frac{(b-t)(c-d)}{(b+1)c}\Big)\frac{d(b-3)\lambda t}{(f+2)rk}\Big] \quad \text{for } c\geqq d>0, \tag{4.7}$$

or if

(4.8) $$a \leqq 2\max\left[\min\left\{1, \frac{b-t-4}{f+2}c\right\}, \frac{(b-t-4)\lambda t}{(f+2)rk}c\right] \quad \textit{for } c>0,\ b>t+4.$$

For simple cases $c=d$, $d=0$ or $c=0$, we have $A_2(c, c; 0)=c$, $A_2(c, 0; 0)=c(b-t-4)/(b-3)$ and $A_2(0, d; 0)=d(t-3)/(b-3)$ respectively. Then, from Theorem 4.1 (a), (b) and (c), we get the better results of Bhattacharya [2] for $T_3(a^*, c^*)$ and $T_5(a^*, c^*)$ in the case of BIBD's. In particular, Theorem 4.1 (a) yields the sufficient conditions:

(4.9) $$a^* \leqq 2\max\left[\min\left\{1, \frac{(r-\lambda)(b-3)}{\lambda t(f+2)}c^*\right\}, \frac{(r-\lambda)(b-3)}{rk(f+2)}c^*\right] \quad \text{for } T_3(a^*, c^*),$$

(4.10) $$a^* \leqq 2\max\left[\min\left\{1, \frac{(r-\lambda)(b-t-4)}{\lambda t(f+2)}c^*\right\}, \frac{(r-\lambda)(b-t-4)}{rk(f+2)}c^*\right] \quad \text{for } T_5(a^*, c^*).$$

These sufficient conditions are obtained by the use of the information that $\rho \geqq \rho_0 = \lambda t/(r-\lambda)$, so that they are better than those of Bhattacharya [2] which are given without using the information on ρ. For the estimator $T_4(a^*, c^*)$ of Bhattacharya [2], it is expressed as a special case of estimators similarly induced from (2.1) and (2.2) by the above match ups without W_j's (i. e. $q \sim 0$). Thus from Theorem 2.3 (a), (b) and (c), we also get the better results of Bhattacharya [2] for $T_4(a^*, c^*)$ in the case of BIBD's. For instance, its sufficient condition is given by

(4.11) $$a^* \leqq 2\max\left[\min\left\{1, \frac{(r-\lambda)(b-t-1)}{\lambda t(f+2)}c^*\right\}, \frac{(r-\lambda)(b-t-1)}{rk(f+2)}c^*\right] \quad \text{for } T_4(a^*, c^*).$$

While our scope is limited to BIBD's, it should be noted that these results for $T_3(a^*, c^*)$, $T_4(a^*, c^*)$ and $T_5(a^*, c^*)$ are extended to any incomplete block design as is shown by Bhattacharya [2].

Now, using Theorem 4.1, we shall find a sufficient condition for Yates' estimate, which is still the most widely used, being better than the intrablock estimate. First we shall get a sufficient condition for the nontruncated Yates' estimate given by Graybill and Weeks [8] as

(4.12) $$\hat{\tau}_Y = x_1 + \frac{(r-\lambda)S^2/f}{\lambda t(S^2/f + k\hat{\sigma}_\beta^2) + (r-\lambda)S^2/f}(y_1 - x_1),$$

where

(4.13) $$\hat{\sigma}_\beta^2 = \frac{1}{t(r-1)}\left\{\frac{\lambda t(r-\lambda)}{rk^2}(\boldsymbol{x}-\boldsymbol{y})'(\boldsymbol{x}-\boldsymbol{y}) + S^{*2} - \frac{b-1}{f}S^2\right\}.$$

Note that the relations $r(k-1)=\lambda(t-1)$ and $bk=rt$ hold for BIBD's. Rearranging and calculating the coefficients of S^2, S^{*2} and $(\boldsymbol{x}-\boldsymbol{y})'(\boldsymbol{x}-\boldsymbol{y})$ in (4.12) and (4.13), we can see that the estimate (4.12) is represented as $\hat{\tau}(a, c, d)$ in (4.6) with $a=(r-1)/(r-k)$, $c=f/(b-t)$ and $d=\lambda tf/\{rk(b-t)\}$. Since c is not equal to d, we can not employ sufficient conditions given by Brown and Cohen [4], Khatri and Shah [9] and Bhattacharya [2]. But it is noted that the condition $c>d$ in Theorem 4.1 is satisfied because $rk-\lambda t=r-\lambda>0$. Hence we can use Theorem 4.1 (d) and get a simple sufficient condition

$$(4.14)\qquad \frac{r-1}{r-k}\leqq 2\max\left[\min\left\{1, \left(1+\frac{(b-t)(r-\lambda)}{(b+1)rk}\right)\frac{\lambda tf(b-3)}{rk(b-t)(f+2)}\right\},\right.$$
$$\left.\left(1+\frac{(b-t)(r-\lambda)}{(b+1)rk}\right)\left(\frac{\lambda t}{rk}\right)^2\frac{f(b-3)}{(b-t)(f+2)}\right].$$

Calculating values of both sides in (4.14) for all asymmetrical BIBD's listed in Fisher-Yates table [6], we can see that this inequality holds, i. e. $\hat{\tau}_Y$ offers uniform improvement over x_1, except for a design $r=3$, $t=4$, $b=6$, $k=2$, $\lambda=1$. This design is one of two exceptional designs in Bhattacharya [1], and he proved that $\hat{\tau}_Y$ does not have the desired property. For the other exceptional design: $r=4$, $t=5$, $b=10$, $k=2$, $\lambda=1$, we can conclude by our sufficient condition (4.14) that $\hat{\tau}_Y$ is better than x_1. Note that the well-known Yates' estimate $\hat{\tau}_Y^*$ is the truncated form given as

$$(4.15)\qquad \begin{aligned}\hat{\tau}_Y^* &= \hat{\tau}_Y && \text{if } \hat{\sigma}_\beta^2>0\,,\\ &= x_1+(r-\lambda)(y_1-x_1)/(rk) && \text{if } \hat{\sigma}_\beta^2\leqq 0\,,\end{aligned}$$

and that it is superior to the untruncated estimate $\hat{\tau}_Y$ as is shown in Seshadri [10]. It follows that Yates' estimate is better than the intrablock estimate for all asymmetrical BIBD's listed in Fisher-Yates table [6] except for the design $r=3$, $t=4$, $b=6$, $k=2$, $\lambda=1$.

4.3. Estimation of common regression coefficients of two regression models. Let

$$(4.16)\qquad \boldsymbol{y}_i=\boldsymbol{X}_i\boldsymbol{\beta}+\boldsymbol{\varepsilon}_i\,,\qquad i=1, 2\,,$$

be two regression models with common regression coefficients where $\boldsymbol{y}_i$ is a $n_i\times 1$ vector of observations, $\boldsymbol{X}_i$ is a known $n_i\times p$ matrix of rank p, $\boldsymbol{\beta}$ is a $p\times 1$ vector of unknown parameters and $\boldsymbol{\varepsilon}_i$ is a $n_i\times 1$ vector of errors having p-variate normal distribution $N_p(\boldsymbol{0}, \sigma_i^2\boldsymbol{I}_{n_i})$, $i=1, 2$. The least square estimator $\hat{\boldsymbol{\beta}}_i=(\boldsymbol{X}_i'\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i'\boldsymbol{y}_i$ has $N_p(\boldsymbol{\beta}, \sigma_i^2(\boldsymbol{X}_i'\boldsymbol{X}_i)^{-1})$, and the residual sum of squares $S_i=(\boldsymbol{y}_i-\boldsymbol{X}_i\hat{\boldsymbol{\beta}}_i)'(\boldsymbol{y}_i-\boldsymbol{X}_i\hat{\boldsymbol{\beta}}_i)$ has $\sigma_i^2\chi_{n_i-p}^2$-distribution $(n_i>p)$, $i=1, 2$.

To estimate common coefficients $\boldsymbol{\beta}$, we consider combined estimators of the form

$$\hat{\boldsymbol{\beta}}=\hat{\boldsymbol{\beta}}_1+\boldsymbol{\phi}\cdot(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)\,, \tag{4.17}$$

where

$$\boldsymbol{\phi}=aS_1[S_1\boldsymbol{X}_2'\boldsymbol{X}_2+\boldsymbol{X}_1'\boldsymbol{X}_1\{cS_2+d(\hat{\boldsymbol{\beta}}_1-\hat{\boldsymbol{\beta}}_2)'\boldsymbol{X}_2'\boldsymbol{X}_2(\hat{\boldsymbol{\beta}}_1-\hat{\boldsymbol{\beta}}_2)\}]^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2 \tag{4.18}$$

with nonnegative constants a, c and d $(c+d>0)$ suitably chosen. These estimators are unbiased and are a special case of Swamy and Mehta [11], whose estimators correspond to (4.17) by interchanging subscripts 1 and 2. The problem is to find a better estimator within the class of (4.17) than the least square estimator $\hat{\boldsymbol{\beta}}_1$ based on the first model only, where the preference of estimators is judged by usual partial ordering between covariance matrices. From the method used by Swamy and Mehta [11] and Theorem 2.3, we get

THEOREM 4.2. *Let $n_i\geqq p+1$ for $i=1, 2$. Suppose that one of the next three conditions holds:* (i) *$n_2>2$ if c, $d>0$,* (ii) *$n_2>p+4$ if $d=0$ or* (iii) *$p>2$ if $c=0$. Put $A_3(c, d\,;0)=E[\{c(1-Z_3)+dZ_3\}^{-1}]/E[\{c(1-Z_3)+dZ_3\}^{-2}]$ where Z_3 has a beta distribution with parameters $((p+2)/2, (n_2-p)/2)$. Assume that*

(C-1) $c\geqq d\geqq 0$ *or* $c=0$ $(c+d>0)$,

(C-2) $2\min\limits_{1\leq i\leq p}(\lambda_i)\geqq\max\limits_{1\leq i\leq p}(\lambda_i)$, *for eigen values* $\lambda_1, \cdots, \lambda_p$ *of* $(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2$.

However, we need not assume the condition (C-2) *when $c=d$ or $d=0$. Then*

(a) *$\hat{\boldsymbol{\beta}}$ is better than $\hat{\boldsymbol{\beta}}_1$ if $a\leqq 2\min\{1, A_3(c, d\,;0)(n_2-2)/(n_1-p+2)\}$.*

(b) *Given $A_3(c, d\,;0)(n_2-2)/(n_1-p+2)\leqq 1$, $\hat{\boldsymbol{\beta}}$ is better than $\hat{\boldsymbol{\beta}}_1$ if and only if $a\leqq 2A_3(c, d\,;0)(n_2-2)/(n_1-p+2)$.*

(c) *Given $a\leqq 2$, $\hat{\boldsymbol{\beta}}$ is better than $\hat{\boldsymbol{\beta}}_1$ if and only if $A_3(c, d\,;0)\geqq a(n_1-p+2)/\{2(n_2-2)\}$.*

(d) *$\hat{\boldsymbol{\beta}}$ is better than $\hat{\boldsymbol{\beta}}_1$ if*

$$a\leqq 2\min\left\{1, \frac{(n_2-2)d}{n_1-p+2}\left(1+\frac{(n_2-p)(c-d)}{(n_2+2)c}\right)\right\}\qquad \text{for } c\geqq d>0\,, \tag{4.19}$$

or if

$$a\leqq 2\min\left\{1, \frac{n_2-p-4}{n_1-p+2}c\right\}\qquad \text{for } c>0,\ n_2>p+4\,. \tag{4.20}$$

PROOF. First we write the covariance matrix of $\hat{\boldsymbol{\beta}}$ as

$$\begin{aligned} Cov(\hat{\boldsymbol{\beta}})=Cov(\hat{\boldsymbol{\beta}}_1)+E[&\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\phi}'+\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})' \\ &+(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\phi}']\,, \end{aligned} \tag{4.21}$$

so that $Cov(\hat{\boldsymbol{\beta}}_1)-Cov(\hat{\boldsymbol{\beta}})$ is *psd* if, and only if,

$$E[\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_1-\hat{\boldsymbol{\beta}}_2)(\hat{\boldsymbol{\beta}}_1-\hat{\boldsymbol{\beta}}_2)'\boldsymbol{\phi}'+\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})'+(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\phi}']\leqq 0. \tag{4.22}$$

To diagonalize the matrix (4.22), we consider a nonsingular matrix $\boldsymbol{Q}=(\boldsymbol{q}_1, \cdots, \boldsymbol{q}_p)$ such that $\boldsymbol{X}_1'\boldsymbol{X}_1=\boldsymbol{Q}\boldsymbol{Q}'$ and $\boldsymbol{X}_2'\boldsymbol{X}_2=\boldsymbol{Q}\boldsymbol{D}_\lambda\boldsymbol{Q}'$ where $\boldsymbol{D}_\lambda=\text{diag}(\lambda_1, \cdots, \lambda_p)$ and λ_i's are the eigen values of $(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2$. Then it can be seen that the weighting matrix $\boldsymbol{\phi}$ is diagonalizable because

$$\begin{aligned}\boldsymbol{Q}'\boldsymbol{\phi}\boldsymbol{Q}'^{-1}&=aS_1[S_1\boldsymbol{D}_\lambda+\boldsymbol{I}_p\{cS_2+d\textstyle\sum_{j=1}^{p}\lambda_j(\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2\}]^{-1}\boldsymbol{D}_\lambda\\&=\text{diag}\Big(\frac{aS_1}{S_1+(c/\lambda_i)S_2+(d/\lambda_i)\sum_{j=1}^{p}\lambda_j(\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2}\Big)\\&=\text{diag}(\phi_1, \cdots, \phi_p) \qquad \text{(say).}\end{aligned} \tag{4.23}$$

Note that $\boldsymbol{Q}'\hat{\boldsymbol{\beta}}_1$ has p-variate normal distribution $N_p(\boldsymbol{Q}'\boldsymbol{\beta}, \sigma_1^2\boldsymbol{I}_p)$ and that $\boldsymbol{Q}'\hat{\boldsymbol{\beta}}_2$ has $N_p(\boldsymbol{Q}'\boldsymbol{\beta}, \sigma_2^2\boldsymbol{D}_\lambda^{-1})$. Then we can multiply the matrix (4.22) on the left by $\boldsymbol{Q}'$ and on the right by $\boldsymbol{Q}$, so that we have

$$\begin{aligned}&\boldsymbol{Q}'E[\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\phi}'+\boldsymbol{\phi}(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})'+(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)'\boldsymbol{\phi}']\boldsymbol{Q}\\&=E[\text{diag}(\phi_i)E[\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)\}\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)\}' \mid (\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2, j\geqq 1]\text{diag}(\phi_i)\\&\quad+\text{diag}(\phi_i)E[\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)\}\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})\}' \mid (\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2, j\geqq 1]\\&\quad+E[\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_1-\boldsymbol{\beta})\}\{\boldsymbol{Q}'(\hat{\boldsymbol{\beta}}_2-\hat{\boldsymbol{\beta}}_1)\}' \mid (\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2, j\geqq 1]\text{diag}(\phi_i)]\\&=\text{diag}(E[\phi_i^2(\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_2-\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_1)^2+2\phi_i(\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_2-\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_1)(\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_i'\boldsymbol{\beta})]).\end{aligned} \tag{4.24}$$

Using the same method of Brown and Cohen [4] and Khatri and Shah [9], each diagonal element in the last expression of (4.24) is written as

$$\sigma_1^2E[(1+\rho_i)\bar{\phi}_i^2-2\bar{\phi}_i], \qquad i=1, \cdots, p, \tag{4.25}$$

where $\rho_i=\sigma_2^2/(\lambda_i\sigma_1^2)>0$ and $\bar{\phi}_i$ is the same as ϕ_i in (4.23) except that $(\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_2-\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_1)^2$ is replaced by a random variable having $(\sigma_1^2+\sigma_2^2/\lambda_i)\chi_3^2$-distribution. Hence $Cov(\hat{\boldsymbol{\beta}}_1)-Cov(\hat{\boldsymbol{\beta}})$ is psd if, and only if,

$$a\leqq 2\inf_{\rho_i>0}\Big\{\frac{E[r_i]}{E[r_i^2]}\Big\}, \qquad i=1, \cdots, p, \tag{4.26}$$

where $r_i=(1+\rho_i)\bar{\phi}_i/a$. Let us make the following match ups ($\sim$) with the terms used in Section 2: for each i, $X\sim\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_1$, $Y\sim\boldsymbol{q}_i'\hat{\boldsymbol{\beta}}_2$, $\rho\sim\rho_i$, $\rho_0\sim 0$, $r(\rho)\sim r_i$, $m_1\sim n_1-p$, $m_2\sim n_2-p$, $q\sim p-1$, $\alpha_0\sim 1$, $\beta_0\sim 1/\lambda_i$; $\alpha_j\sim 1$, $\beta_j\sim 1/\lambda_j$, $W_j\sim(\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_j'\hat{\boldsymbol{\beta}}_2)^2$, $j=1, \cdots, i-1$; $\alpha_j\sim 1$, $\beta_j\sim 1/\lambda_{j+1}$, $W_j\sim(\boldsymbol{q}_{j+1}'\hat{\boldsymbol{\beta}}_1-\boldsymbol{q}_{j+1}'\hat{\boldsymbol{\beta}}_2)^2$, $j=i, \cdots, p-1$. Then, we can employ Theorem 2.3 to obtain the conditions in Theorem 4.2, which complete the proof.

For special cases $c=d$, $d=0$ or $c=0$, we have $A_3(c, c; 0)=c$, $A_3(c, 0; 0)=c(n_2-p-4)/(n_2-2)$ or $A_3(0, d; 0)=d(p-2)/(n_2-2)$ respectively. Then Theorem

4.2 (c) gives that the necessary and sufficient conditions for $\hat{\beta}$ being better than $\hat{\beta}_1$ are $a \leqq 2c(n_2-2)/(n_1-p+2)$ for $0<a\leqq 2$ and $c=d$; $a\leqq 2c(n_2-p-4)/(n_1-p+2)$ for $0<a\leqq 2$ and $d=0$. From these conditions, we can easily see that $\hat{\beta}$ is better than $\hat{\beta}_1$ if $1\leqq 2c(n_2-2)/(n_1-p+2)$ for $0<a\leqq 1$ and $c=d$, or if $1\leqq 2c(n_2-p-4)/(n_1-p+2)$ for $0<a\leqq 1$ and $d=0$, which was given by Swamy and Mehta [11]. They claimed wrongly these conditions to be necessary and sufficient.

Acknowledgements. The author is extremely thankful to Professor N. Sugiura for his helpful criticisms and valuable advice during the progress of this work. Especially, the proof of Theorem 2.3 is owing to his suggestion. The author is also thankful to the referee for his careful reading of the original manuscript. His comments led to improvements in the presentation of the results.

References

[1] Bhattacharya, C. G., Yates type estimators of a common mean. Ann. Inst. Statist. Math. **30** (1978), 407-414.
[2] ———, Estimation of a common mean and recovery of interblock information. Ann. Statist. **8** (1980), 205-211.
[3] ———, Two inequalities with an application. Ann. Inst. Statist. Math. **36** (1984), 129-134.
[4] Brown, L. D. and Cohen, A., Point and confidence estimation of a common mean and recovery of interblock information. Ann. Statist. **2** (1974), 963-976.
[5] Exton, H., Handbook of Hypergeometric Integrals. John Wiley & Sons, 1978.
[6] Fisher, R. A. and Yates, F., Statistical Tables for Biological, Agricultural and Medical Research. 6th Ed. Oliver and Boyd, Edinburgh, 1963.
[7] Graybill, F. A. and Deal, R. B., Combining unbiased estimators. Biometrics **15** (1959), 543-550.
[8] ——— and Weeks, D. L., Combining inter-block and intra-block information in balanced incomplete blocks. Ann. Math. Statist. **30** (1959), 799-805.
[9] Khatri, C. G. and Shah, K. R., Estimation of location parameters from two linear models under normality. Commun. Statist. **3** (1974), 647-663.
[10] Seshadri, V., Comparison of combined estimators in balanced incomplete blocks. Ann. Math. Statist. **37** (1966), 1832-1835.
[11] Swamy, P. A. V. B. and Mehta, J. S., Estimation of common coefficients in two regression equations. J. Econometrics **10** (1979), 1-14.
[12] Yates, F., The recovery of inter-block information in balanced incomplete block designs. Ann. of Eugenics **10** (1940), 317-325.

Doctoral Program in Mathematics
University of Tsukuba
Sakuramura, Ibaraki 305
Japan