

ESTIMATION OF A COMMON MEAN AND RECOVERY OF INTERBLOCK INFORMATION

BY C. G. BHATTACHARYA

University of Waterloo

Consider the problem of combining two unbiased estimators of a parameter when the estimators are known to be independent normal variables with unknown variances possibly unequal. The two one parameter families of estimators studied in Brown and Cohen, and Khatri and Shah, are accommodated in a single two parameter family studied in this paper and the results in the two papers are unified. For the type of estimators considered by Brown and Cohen, this paper not only offers a generalization but also a significant improvement. This improvement concerns the main result in Theorem 2.1 of Brown and Cohen and has bearing on their entire paper except the last section on interval estimation. Extensions of Brown and Cohen's Theorem 4.1 concerning the point estimation of the common mean of K -populations and Theorem 5.1 concerning interval estimation of the common mean of two populations are also presented.

1. Introduction. The problem of estimating the common mean of two normal distributions and the related problem of recovery of interblock information have been studied in several papers. Since Yates (1939, 1940) initiated the whole subject and his work was extended by Nair (1944) and Rao (1947, 1956) many have turned their attention to the problem of constructing a combined unbiased estimator with a variance uniformly smaller than that of the intrablock estimator (first of the two sample means in the common mean problem). Of these, the earlier works by Graybill and Deal (1959), Seshadri (1963 a, b), Shah (1964) and Stein (1966) are applicable only to some special designs. Recent works by Brown and Cohen (1974) and Khatri and Shah (1974), with which this paper is really concerned, are applicable to any incomplete block design with more than 3 blocks. For the sake of brevity, these two papers will henceforth be referred to as BC and KS, respectively. For the same reason a statement like ' T is uniformly better than x ' will be written as ' T ub x '.

Both BC and KS, in fact considered a whole family of estimators depending on a single parameter and while the former required an upper bound on their parameter, the latter required a lower bound. The two families of estimators mentioned above can be treated as particular cases of a two parameter family and the desired property could be guaranteed by a single condition on the two parameters. The upper bound set on their parameter by BC was somewhat crude and involved a complicated expression which called for a table given in their paper. Thus, the

Received October 1975; revised October 1978.

AMS 1970 subject classifications. Primary 62F10; secondary 62K10, 62K15.

Key words and phrases. Common mean, unbiased estimators, connected binary equireplicate incomplete block designs, balanced incomplete block designs, interblock information, confidence intervals.

present paper unifies these two classes of estimators and improves the result of BC for the subclass which corresponds to their estimator. The improved upper bound of the BC parameter, as given in this paper, is not only more precise but also is a simple expression, for which no table is required. For the estimator considered by BC, the improved upper bound of their parameter is, in fact, shown to be the best possible. Incidentally, it was noticed that Theorem 4.1 and Theorem 5.1 of BC can be easily extended to the corresponding two parameter family and these have been presented in Remark 2.4 after the main result in Section 2. Finally, Section 3 contains application of the results of Section 2 to the two main problems, namely, estimation of a common mean and recovery of inter-block information, which motivated most of the studies in this field.

2. Main results. Let $x, y, S, T, W_i, i = 1, 2, \dots, q$ be independent random variables where $x \sim N(\mu, \alpha_0\sigma^2), y \sim N(\mu, \beta_0\eta^2), S/\sigma^2 \sim \chi_m^2, T/\eta^2 \sim \chi_n^2,$ and $W_i/(\alpha_i\sigma^2 + \beta_i\eta^2) \sim \chi_1^2, i = 1, 2, \dots, q.$ Let

$$(2.1) \quad \hat{\mu} = x + \phi(y - x)$$

with

$$(2.2) \quad \phi = aS/[S + c\{T + (y - x)^2/\beta_0 + \sum_{i=1}^q W_i/\beta_i\}]$$

and a and c are constants to be suitably chosen.

Let W_0 be independent of $S, T,$ and $W_i, i = 1, 2, \dots, q$ such that $W_0/(\alpha_0\sigma^2 + \beta_0\eta^2) \sim \chi_3^2$ and let ϕ^* be the expression obtained by replacing $(y - x)^2$ in ϕ by $W_0.$ Both BC and KS have shown that $V(\hat{\mu}) \leq V(x)$ for all values of τ iff

$$(2.3) \quad (1 + \tau)E(\phi^{*2}) \leq 2E\phi^* \quad \text{for every } \tau$$

where $\tau = \beta_0\eta^2/(\alpha_0\sigma^2).$ Let $r = (1 + \tau)\phi^*/a.$ It is easy to verify that (2.3) is equivalent to $a \leq 2E(r)/E(r^2)$ for every $\tau.$ Thus,

THEOREM 2.1. $\hat{\mu}$ is uniformly better than x iff $a \leq 2\theta$ where $\theta = \inf_{\tau} E(r)/E(r^2).$

Since it is not easy to evaluate θ in all cases attempts will be made to obtain some nontrivial lower bounds for $\theta.$ Let

$Z_1 = S/\sigma^2, Z_2 = T/\eta^2 + \sum_{i=0}^q W_i/(\alpha_i\sigma^2 + \beta_i\eta^2), u_j = W_j/[(\alpha_j\sigma^2 + \beta_j\eta^2)Z_2],$
 $j = 0, 1, \dots, q$ and let $u = \beta_0(\sum \alpha_j u_j/\beta_j)/\alpha_0.$ It is easy to see that $Z_1 \sim \chi_m^2,$
 $Z_2 \sim \chi_{n+q+3}^2,$ and that u, Z_1, Z_2 are all independently distributed. It is also not difficult to see that r can be written in the form

$$(2.4) \quad r = Z_1/[(1 - \gamma)Z_1 + dZ_2h(u, \gamma)]$$

where $d = c\alpha_0/\beta_0, \gamma = \tau/(1 + \tau), h(u, \gamma) = u(1 - \gamma) + \gamma.$ Let $\delta = \inf_{u, \gamma} f(u, \gamma)$ where $f(u, \gamma) = E(r|u)/E(r^2|u).$ Clearly, $\delta \leq \theta$ and hence it would be sufficient to have $a \leq 2\delta$ to ensure $v(\hat{\mu}) \leq V(x).$ Let primes denote derivations with respect to $\gamma.$ Then direct computation from (2.4) shows that

$$(2.5) \quad r' = [r^2 - (1 - u)r]/h(u, \gamma).$$

Also $f' = -g(u)/[h(u, \gamma)E^2(r^2|u)]$ where

$$g(u) = 2E(r|u)E(r^3|u) - (1 - u)E(r|u)E(r^2|u) - E^2(r^2|u).$$

The following lemma can now be proved easily.

LEMMA 2.1. *If $E(r'|u) \geq 0$, then $f' \leq 0$.*

PROOF. If $E(r'|u) \geq 0$, then $E(r^2|u) \geq (1 - u)E(r|u)$ and hence $g(u) \geq 2E(r|u)E(r^3|u) - 2E^2(r^2|u) \geq 0$ and hence $f' \leq 0$. To evaluate δ , first consider the case where $u \geq 1$. Here $r' \geq 0$ and hence $E(r'|u) \geq 0$ for $0 \leq \gamma < 1$. Thus, $f' \leq 0$ for $u \geq 1$. This gives $\inf_{\gamma} f(u, \gamma) = \lim_{\gamma \rightarrow 1} f(u, \gamma) = da_0$ where $a_0 = (n + q - 1)/(m + 2)$.

Since $r'' \geq 0$ for $u > 0$, it follows that r' is a nondecreasing function of γ and hence either (i) $E(r'|u) \geq 0$ for $\gamma \in (0, 1)$ or (ii) $E(r'|u) < 0$ for $\gamma \in (0, 1)$ or (iii) $E(r'|u) < 0$ for $0 < \gamma < \lambda(u)$ and $E(r'|u) \geq 0$ for $\lambda(u) \leq \gamma < 1$.

Thus for $0 < u < 1$, if $E(r'|u) \geq 0$, then $\inf_{\gamma} f(u, \gamma) = \lim_{\gamma \rightarrow 1} f(u, \gamma) = da_0$. On the other hand if for $0 < u < 1$, $E(r'|u) < 0$, then (2.5) gives $f(u, \gamma) = E(r|u)/E(r^2|u) > 1/(1 - u)$. Thus, $\delta = \inf_{\gamma} f(u, \gamma) = da_0$ if $da_0 \leq 1$ and $\delta \in (1, da_0)$ otherwise. For $da_0 \leq 1$, $da_0 = \delta \leq \theta$ and hence $da_0 = \theta$. In view of Theorem 2.1 we have thus proved

THEOREM 2.2. *If $n + q \geq 2$ and $a_0 = (n + q - 1)/(m + 2)$, then*

- (a) *for $a \leq 2 \min[1, da_0]$, $\hat{\mu}$ is uniformly better than x ;*
- (b) *if $da_0 \leq 1$, $\hat{\mu}$ is uniformly better than x iff $a \leq 2da_0$;*
- (c) *for fixed $a \leq 2$, $\hat{\mu}$ is uniformly better than x iff $d \geq a/(2a_0)$.*

REMARK 2.1. Taking $a = 1$, Theorem 2.2(c) gives the following result proved by KS in a different way: if $n + q \geq 2$ and $a_0 = (n + q - 1)/(m + 2)$, then the estimator $\hat{\mu}$ with $a = 1$ is uniformly better than x iff $d \geq (m + 2)/[2(n + q - 1)]$.

The following useful modifications of $\hat{\mu}$ can be dealt in the same way as $\hat{\mu}$. Let $\hat{\mu}_i = x + \phi_i(y - x)$, $i = 1, 2, 3$ where

$$\begin{aligned} \phi_1 &= aS/[S + c\{T + (y - x)^2/\beta_0\}] \\ \phi_2 &= aS/[S + c\{(y - x)^2/\beta_0 + \sum_{j=1}^k W_j/\beta_j\}] \\ \phi_3 &= aS/[S + cT]. \end{aligned}$$

Thus,

THEOREM 2.3. *Theorem 2.2 holds word by word for each $\hat{\mu}_i$ provided the expression for a_0 in that theorem is replaced by*

$$a_0 = f_i/(m + 2) \quad \text{where } f_1 = n - 1, \quad f_2 = q - 1, \quad f_3 = n - 4.$$

REMARK 2.2. The estimators T_a and $T_a(1)$ of BC are particular cases of $\hat{\mu}_1$ and $\hat{\mu}_3$, respectively and application of Theorem 2.3 to these estimators leads to improvement of Theorem 2.1 and Theorem 2.2 of their paper. For details refer to Section 3 on application.

REMARK 2.3. Taking $a = 1$, Theorem 2.3 yields the following:

$$\hat{\mu}_3 \text{ ub } x \text{ iff } d \geq (m+2)/2(n-4), \quad \text{provided } n \geq 5.$$

In view of symmetry considerations, it follows that, $\hat{\mu}_3$ is uniformly better than both x and y iff $(m+2)/[2(n-4)] \leq d \leq 2(m-4)/(n+2)$, which is possible whenever $(m-6)/(n-6) \geq 16$. This result, which is an improvement of a similar result by Graybill and Deal (1959), was obtained by KS in a quite different way.

REMARK 2.4. As $\hat{\mu}_3$ generalizes $T_a(1)$ of BC the following generalize Theorem 4.1 and Theorem 5.1 of their paper:

(a) Let $x_i, S_i, i = 1, 2, \dots, k$ be independent random variables such that $x_i \sim N(\mu, \delta_i \sigma_i^2)$ and $S_i/\sigma_i^2 \sim \chi^2(m_i)$. Let $b_i, c_i, i = 1, 2, \dots, k-1$ be arbitrary sequences of positive numbers such that

$$(2.6) \quad b_i < \min[1, 2d_i(m_{i+1}-4)/(m_i+2)], \quad \text{where } d_i = c_i \delta_i / \delta_{i+1}.$$

Let $a_1 = b_1$ and $a_i = b_i(1 - \sum_{j=1}^{i-1} a_j)$, $i = 2, 3, \dots, k-1$. Then $\hat{\mu}_3(k) = x_1 + \sum_{i=1}^{k-1} (x_{i+1} - x_i) a_i S_1 / (S_1 + c_i S_{i+1})$, is uniformly better than x_1 , provided $m_i \geq 5$, $i = 2, 3, \dots, k$. The result can be proved in the same way as in BC once it is noted that condition (2.6) ensures that for each i , $b_i < 1$ and the estimator $l_i = x_1 + (x_{i+1} - x_1) b_i / (S_1 + c_i S_{i+1}) \text{ub } x_1$.

(b) Consider the random intervals

$$(2.7) \quad x \pm t_m(\alpha)(\alpha_0 S/m)^{\frac{1}{2}} \dots \text{ and,}$$

$$(2.8) \quad \hat{\mu}_3 \pm t_m(\alpha)(\alpha_0 S/m)^{\frac{1}{2}} \dots$$

where $t_m(\alpha)$ stands for the two-tailed α critical value determined from Student's t -distribution with m degrees of freedom and the rest of the symbols are as defined in the beginning of this section. Then there exists an $\epsilon(\alpha) \in (0, 1)$ such that for $a \leq \epsilon(\alpha)$ the confidence interval (2.8) is better than the confidence interval (2.7) in the sense that both intervals have the same length and the probability of coverage of (2.8) is uniformly greater than that of (2.7), provided $n \geq 5$.

The proof of the result is analogous to that in BC and is omitted.

3. Applications.

(a) *Estimation of a common mean.* Let (x_1, x_2, \dots, x_m) and (y_1, y_2, \dots, y_m) be independent random samples from two normal populations having a common unknown mean and unknown variances σ_x^2 and σ_y^2 respectively. Consider the problem of estimating μ . Let $\bar{x} = \sum x_i/m$, $s_x^2 = \sum (x_i - \bar{x})^2 / \{m(m-1)\}$ and let \bar{y}, s_y^2 be defined similarly. Let,

$$T_1(a, c) = \bar{x} + (\bar{y} - \bar{x}) a s_x^2 / \left[s_x^2 + c \left\{ s_y^2 + (\bar{y} - \bar{x})^2 / (n-1) \right\} \right]$$

$$T_2(a, c) = \bar{x} + (\bar{y} - \bar{x}) a s_x^2 / (s_x^2 + c s_y^2).$$

These include as particular cases the estimators $T_a, T_a(1)$ of BC and μ^*, μ^{**} of KS. In fact,

$$T_a = T_1(a, (n-1)/(n+2)), \quad T_a(1) = T_2(a, 1) \\ \mu^* = T_1(1, c(n-1)/(m-1)), \quad \mu^{**} = T_2(1, c(n-1)/(m-1)).$$

It can be seen that $T_1(a, c)$ and $T_2(a, c)$ are of the same forms as $\hat{\mu}_1$ and $\hat{\mu}_3$, respectively. Note that $d = c(m-1)/(n-1)$ in each case and the values of a_0 are $(n-2)/(m+1)$ and $(n-5)/(m+1)$, respectively. It follows from Theorem 2.3 that $T_1(a, c) \text{ub } \bar{x}$ for all $a \leq 2 \min[1, d(n-2)/(m+1)]$ provided $m \geq 2, n \geq 3$. Similarly $T_2(a, c) \text{ub } \bar{x}$ for $a \leq 2 \min[1, d(n-5)/(m+1)]$, provided $m \geq 2, n \geq 6$. Given $a \leq 2$, as is the case with KS estimators it further follows that $T_1(a, c) \text{ub } \bar{x}$ iff $d \geq (\frac{1}{2})a(m+1)/(n-2)$ and $T_2(a, c) \text{ub } \bar{x}$ iff $d \geq (\frac{1}{2})a(m+1)/(n-5)$. From these, the results concerning KS estimators are immediate. Given c such that $da_0 \leq 1$, as is the case with BC estimators, it also follows that $T_1(a, c) \text{ub } \bar{x}$ iff $a \leq 2d(n-2)/(m+1)$ and $T_2(a, c) \text{ub } \bar{x}$ iff $a \leq 2d(n-5)/(m+1)$. In particular, $T_a \text{ub } \bar{x}$ iff $a \leq 2(m-1)(n-2)/[(m+1)(n+2)]$ and $T_a(1) \text{ub } \bar{x}$ iff $a \leq 2(m-1)(n-5)/\{(m+1)(n-1)\}$. These results are readily seen to be improvements of Theorems 2.1 and 2.2 of BC.

(b) *Recovery of inter-block information.* Consider a connected binary equireplicate incomplete block design. Let $b =$ number of blocks, $k =$ number of plots per block. Let $N(v \times b)$ denote the incidence matrix of the design and let $\text{rank}(N) = t$. Let $\phi_1, \dots, \phi_{v-1}$ denote the $v-1$ characteristic roots of NN' other than rk and assume, without loss of generality that $\phi_1, \dots, \phi_{t-1}$ are the nonzero ones among these. Consider the following canonical reduction of the plot yields, given by Roy and Shah (1962) under an Eisenhart model III (Eisenhart (1947)): $x_i \sim N(\xi_i, a_i\sigma_1^2), i = 1, \dots, v-1; y_i \sim N(\xi_i, b_i\sigma_2^2), i = 1, \dots, t-1; S_1^2/\sigma_1^2 \sim \chi^2(e_1)$ and $S_2^2/\sigma_2^2 \sim \chi^2(e_2)$ where $a_i = k/(rk - \phi_i), b_i = k/\phi_i, e_1 = bk - b - v + 1, e_2 = b - t$ and $\xi_1, \dots, \xi_{v-1}, \sigma_1^2, \sigma_2^2$ are unknown. The statistics $x_1, \dots, x_{v-1}, y_1, \dots, y_{t-1}, S_1^2, S_2^2$ are mutually independent and constitute a set of minimal sufficient statistics for the treatment effects. Consider the problems of estimating ξ_i (a canonical contrast) for some $i \leq t-1$. Let

$$T_3(a, c) = x_i + (y_i - x_i)aS_1^2 / [S_1^2 + c\{S_2^2 + \sum_{j=1}^{t-1}(y_j - x_j)^2/b_j\}] \\ T_4(a, c) = x_i + (y_i - x_i)aS_1^2 / [S_1^2 + c\{S_2^2 + (y_i - x_i)^2/b_i\}] \\ T_5(a, c) = x_i + (y_i - x_i)aS_1^2 / [S_1^2 + cS_2^2].$$

These include as particular cases the estimators $\hat{\xi}_i$ of KS and $\hat{\mu}_a, \hat{\mu}_a(1), \hat{\mu}_a^*$ of BC. In fact, $\hat{\xi}_i = T_3(1, cb_i/a_i)$,

$$\hat{\mu}_a = T_4(a, e_1(b_i/a_i)/(b-t-3)), \quad \hat{\mu}_a(1) = T_5(a, e_1(b_i/a_i)/(b-t)) \\ \hat{\mu}_a^* = T_3(a, e_1(b_i/a_i)/(b+1)).$$

Note that BC actually concerned themselves only with a BIBD. It can be seen that $T_3(a, c), T_4(a, c)$ and $T_5(a, c)$ are of the same forms as $\hat{\mu}, \hat{\mu}_1$ and $\hat{\mu}_3$, respectively.

Note that $d = ca_i/b_i$ in each case and the values of a_0 are $(b-3)/(e_1+2)$, $(b-t-1)/(e_1+2)$ and $(b-t-4)/(e_1+2)$, respectively. It follows from Theorems 2.2 and 2.3 that $T_3(a, c)$ ub x_i for all $a \leq 2 \min[1, d(b-3)/(e_1+2)]$ provided $b \geq 4$, $T_4(a, c)$ ub x_i for all $a \leq 2 \min[1, d(b-t-1)/(e_1+2)]$ provided $b-t \geq 2$, and $T_5(a, c)$ ub x_i for all $a \leq 2 \min[1, d(b-t-4)/(e_1+2)]$ provided $b-t \geq 5$. Given $a \leq 2$, as is the case with the KS estimator, it further follows that $T_3(a, c)$ ub x_i iff $d \geq (\frac{1}{2})a(e_1+2)/(b-3)$, $T_4(a, c)$ ub x_i iff $d \geq (\frac{1}{2})a(e_1+2)/(b-t-1)$ and $T_5(a, c)$ ub x_i iff $d \geq (\frac{1}{2})a(e_1+2)/(b-t-4)$. Results concerning $\hat{\xi}_i$ of KS is immediate from that concerning $T_3(a, c)$ above.

Given c such that $da_0 \leq 1$, as is the case with BC estimators, $T_3(a, c)$ ub x_i iff $a \leq 2d(b-3)/(e_1+2)$, $T_4(a, c)$ ub x_i iff $a \leq 2d(b-t-1)/(e_1+2)$ $T_5(a, c)$ ub x_i iff $a \leq 2d(b-t-4)/(e_1+2)$. In particular, $\hat{\mu}_a$ ub x_i iff $a \leq 2e_1(b-t-1)/\{(e_1+2)(b-t-3)\}$, $\hat{\mu}_a(1)$ ub x_i iff $a \leq 2e_1(b-t-4)/(e_1+2)$ $(b-t)$, and μ_a^* ub x_i iff $a \leq 2e_1(b-3)/\{(e_1+2)(b+1)\}$.

Note that for a BIBD, $v(x_i) \leq v(y_i)$, as shown in BC and hence the estimators considered above are, in fact, uniformly better than both x_i and y_i under the stated conditions. With $t = v$, as is the case for a BIBD, the results above concerning the BC estimators are readily seen to be improvements of those in Section 3 of BC where the knowledge that $\tau = v(y_i)/v(x_i) > 1$ is used to improve the upper limit of a from a_{\max} to a''_{\max} . Note that a''_{\max} is even more difficult to compute. (There is a misprint of the expression for a''_{\max} given there and the correct expression should be $a''_{\max}(\dots) = 2Ev^{-1}/E \max[2/\{v(1+v)\}, 1/v^2]$).

Acknowledgments. The author is grateful to Professor K. R. Shah for his helpful criticisms and valuable advice during the progress of his work. He is also grateful to the University of Waterloo for offering the research facilities.

REFERENCES

- [1] BROWN, L. D. and COHEN, A. (1974). Point and confidence estimation of a common mean and recovery of inter-block information. *Ann. Statist.* 2 963-976.
- [2] EISENHART, C. (1947). The assumptions underlying the analysis of variance. *Biometrics* 3 1-21.
- [3] GRAYBILL, F. A. and DEAL, R. D. (1959). Combining unbiased estimators. *Biometrics* 15 543-550.
- [4] GRAYBILL, F. A. and SESHADRI, V. (1960). On the unbiasedness of Yates method of estimation using inter-block information. *Ann. Math. Statist.* 31 786-787.
- [5] GRAYBILL, F. A. and WEEKS, D. L. (1959). Combining inter-block and intra-block information in balanced incomplete blocks. *Ann. Math. Statist.* 30 799-805.
- [6] KHATRI, C. G. and SHAH, K. R. (1974). Estimation of location parameters from two linear models under normality. *Comm. Statist.* 3 647-663.
- [7] RAO, C. R. (1947). General method of analysis for incomplete block designs. *J. Amer. Statist. Assoc.* 42 541-561.
- [8] RAO, C. R. (1956). On the recovery of inter-block information in varietal trials. *Sankhyā* 17 105-114.
- [9] ROY, J. and SHAH, K. R. (1962). Recovery of inter-block information. *Sankhyā Ser. A and B* 24 269-280.
- [10] SESHADRI, V. (1963a). Constructing uniformly better estimators. *J. Amer. Statist. Assoc.* 58 172-178.
- [11] SESHADRI, V. (1963b). Combining unbiased estimators. *Biometrics* 19 163-169.

- [12] SHAH, K. R. (1964). Use of inter-block information to obtain uniformly better estimators. *Ann. Math. Statist.* **85** 1064–1078.
- [13] SHAH, K. R. (1970). On the loss of information in combined inter and intra-block estimation. *J. Amer. Statist. Assoc.* **65** 1562–1564.
- [14] YATES, F. (1939). The recovery of intra-block information in varietal trial arranged in three dimensional lattices. *Ann. Eugenics* **9** 136–156.
- [15] YATES, F. (1940). The recovery of inter-block information in balanced incomplete block designs. *Ann. Eugenics* **10** 317–325.

C/O M. G. BHATTACHARYA
ASSISTANT DIRECTOR, SDRD, NSSO
GOVERNMENT OF INDIA
25A, SHAKESPEARE SARANI
CALCUTTA-17, INDIA