Sankhyā : The Indian Journal of Statistics
1983, Volume 45, Series B, Pt. 3, pp. 466–470.

NOTES

USE OF MODIFIED ESTIMATOR IN RECOVERY OF INTER-BLOCK INFORMATION

By C. G. BHATTACHARYA
Indian Statistical Institute

SUMMARY. The paper presents some modifications of the combined intra and inter-block estimate, which are better than the original estimates when we have a-priori knowledge of the lower limit of the ratio of inter to intra-block variances. It also gives a comparison of two such modifications between themselves and with the well known modification based on the truncated estimate of the variance ratio, for which the above property appears to be in doubt except in special situations such as that considered in Shah (1971).

1. Introduction

In the analysis of an incomplete block design with recovery of inter-block information, an estimate (say $\hat{\mu}$) of a canonical contrast (estimable from both intra-block and inter-block analysis) is generally obtained as a weighted average of the intra-block and inter-block estimates of that contrast using suitably chosen random weights. The weight (say φ) given to the inter-block estimate can be expressed in the form $\varphi = 1/(1+\hat{\tau})$ and one may regard $\hat{\mu}$ as an analogue of the best linear unbiased combination of the intra-block and the inter-block estimates in which the unknown ratio (say τ) of the variance of the inter-block estimate to that of the intra-block estimate is replaced by $\hat{\tau}$. τ is generally a known multiple (depending on the design) of the ratio (say ρ) of the inter-block error variance (per plot) to the intra-block error variance. Under the infinite models generally used in the literature ρ cannot be less than 1 and accordingly, τ cannot be less than a known quantity (say τ_0). Hence, if $\hat{\tau} < \tau_0$ it is usually recommended that the value of $\hat{\tau}$ be replaced by τ_0 (see Yates, 1939). More generally, we may have a-priori knowledge that $\rho \geqslant \rho_0$ and then τ_0 may be taken to be the value of τ corresponding to ρ_0. Let $\hat{\mu}_*^{(0)}$ denote the modification of $\hat{\mu}$ corresponding to that of $\hat{\tau}$ as proposed above. Two results which provide theoretical support for modifying $\hat{\mu}$ by $\hat{\mu}_*^{(0)}$ (under certain conditions) are :

(1) $\hat{\mu}_*^{(0)}$ is better than the intra-block estimate for all $\rho \leqslant 2$ (see Shah, 1964).

(2) $\hat{\mu}_*^{(0)}$ is better than $\hat{\mu}$ for all $\rho \geqslant \rho_0$ (see Shah, 1971).

The second result was conjectured by Stein (1966). In the two papers by Shah cited above, there are some common restrictions on φ (see (2.2) and (2.3) of Shah, 1971). Shah (1971) requires an additional assumption (see (2.6) of Shah, 1971) which implies that φ is non-negative. It can be seen that well known methods such as the maximum likelihood method and the MINQUE method lead to estimator of ρ which are not of the form assumed in the two papers by Shah. It is also well known that φ based on the MINQUE method need not be non-negative. In fact the assumption of non-negativity of φ may be violated even if one uses the customary Yates-Rao procedure (untruncated). Interested readers may verify this in the case of canonical contrasts belonging to the eigen-space of the smallest non-zero eigenvalue of the association matrix for PBIB design no. R1 in Bose, *et al.* (1954).

Thus, the present paper generalizes Shah's results under milder conditions. However, in the more general situation considered here, the estimator $\hat{\mu}_*^{(0)}$ does not seem to enjoy the property established in Shah (1971). This leads us to consider two natural alternatives which have the desired property, and to compare these two among themselves and with $\hat{\mu}_*^{(0)}$.

2. Results

Let x, y, S, T, w_i, $i = 1, ..., q$ be independent random variables such that $x \sim N(\mu, \alpha_0\sigma^2)$, $y \sim N(\mu, \beta_0\eta^2)$, $S/\sigma^2 \sim \chi^2_m$, $T/\eta^2 \sim \chi^2_n$, $w_i/(\alpha_i\sigma^2+\beta_i\eta^2) \sim \chi^2_1$, $i = 1, ..., q$ where α_i's and β_i's are known constants and μ, σ^2, η^2 are unknown parameters. Let $w_0 = (y-x)^2$. Interprete x, y, S, T, w_i's as follows: x and y as the intra-block and inter-block estimators of a given canonical contrast which is estimable from both intra-block and inter-block analysis, S and T as the intra-block and inter-block error sum of squares, and w_i's as the squared differences between the inter-block and intra-block estimators of other canonical contrasts which are estimable from both intra-block and inter-block analyses. See Pearce *et al.* (1974) in this connection. We shall use the following theorem which is essentially due to Khatri and Shah (1974).

Theorem 2.1 : *Let $\hat{\mu} = x+\varphi(y-x)$, where φ is a measurable function of S, T, $w_0, ..., w_q$ such that $E\hat{\mu}$ exists. Then,*

$$V(\hat{\mu}) = \alpha_0\sigma^2[1+\bar{E}h(\phi)] \quad \text{... (2.1)}$$

where $h(t) = t^2(1+\tau)-2t$, $\tau = \beta_0\eta^2/\alpha_0\sigma^2$, $\bar{E}$ stands for the expectation with respect to the density w_0f/Ew_0 and f stands for the joint density of $S, T, w_0, ..., w_q$.

In addition, the following lemma will help to make the results transparent.

Lemma 2.1 : *Let $h(t) = t^2(1+\tau)-2t$ where $\tau \geqslant 0$ and let $c = 1/(1+\tau_0)$ where $\tau_0 \geqslant 0$. Then,*

(a) *for every $t < 0$ and $u \in (t, |t|]$,*
$h(t) > h(u)$

(b) *for every $t > c$ and $u \in [c, t)$*
$h(t) > h(u)$ *if* $\tau \geqslant \tau_0$

(c) (i) *for every $t \in (0, c]$,*
$h(t) < 0$ *if* $\tau < 1+2\tau_0$

(ii) $h(c) > 0$ *if* $\tau > 1+2\tau_0$

(d) *for every $t \in (0, c)$*

(i) $h(t) < h(c)$ *if* $\tau \geqslant 1+2\tau_0$

(ii) $h(t) > h(c)$ *if* $\tau \leqslant \tau_0$.

Proof: The proof is elementary and hence omitted.

We first prove

Theorem 2.2: *Let φ be as in Theorem 2.1. Then $\varphi \in [0, 1/(1+\tau_0)]$ a.s. $\Longrightarrow V(\hat{\mu}) \leqslant V(x)$ for all $\tau < 1+2\tau_0$ with strict inequality holding unless $\hat{\mu} = x$ a.s.*

Proof: In view of Theorem 2.1, it suffices to show that for all $\tau < 1+2\tau_0$, $h(\varphi) < h(0) = 0$, if $\varphi \in (0, 1/(1+\tau_0)]$. But this holds by Lemma 2.1(c).

Note that, $\rho_0 = 1$ and $\rho \leqslant 2 \Longrightarrow \tau \leqslant 2\tau_0 \Longrightarrow \tau \leqslant 1+2\tau_0$. Hence Theorem 2.2 yields the result of Shah (1964). We now prove

Theorem 2.3: *Let φ be as in Theorem 2.1. Let $\lambda(\varphi)$ be a measurable function of φ such that $\lambda(\varphi) \in [0, \min\{|\varphi|, 1/(1+\tau_0)\}]$ if $\varphi < 0$. Let $\hat{\mu}_* = x+\varphi_*(y-x)$ where*

$$\begin{aligned} \varphi_* &= \min[\varphi, 1/(1+\tau_0)] \quad && \text{if } \varphi \geqslant 0 \\ &= \lambda(\varphi) && \text{otherwise.} \end{aligned}$$

Then

(i) *$V(\hat{\mu}_*) \leqslant V(\hat{\mu})$ for all $\tau \geqslant \tau_0$, with strict inequality holding unless $\hat{\mu}_* = \hat{\mu}$ a.s.*

(ii) *$V(\hat{\mu}_*) \leqslant V(x)$ for all $\tau < 1+2\tau_0$, with strict inequality holding unless $\hat{\mu}_* = x$ a.s.*

(iii) *$V(\hat{\mu}) \leqslant V(x)$ for all $\tau \geqslant 1+2\tau_0 \Longrightarrow V(\hat{\mu}_*) \leqslant V(x)$ for all $\tau \geqslant 0$.*

Proof: (i) It suffices to show that $h(\varphi) > h(\varphi_*)$ if $\varphi_* \neq \varphi$. But this follows from Lemma 2.1(a) and (b) since $\varphi_* \in [\phi, |\varphi|]$ if $\varphi < 0$, $\varphi_* = 1/(1+\tau_0)$ if $\varphi > 1/(1+\tau_0)$, $\varphi_* = \lambda$ otherwise.

(ii) follows from Theorem 2.2,

(iii) follows from (i) and (ii).

Let $\hat{\mu}, \varphi$ be as in Theorem 2.1 and let

$$\hat{\mu}_*^{(i)} = x+\varphi_*^{(i)}(y-x), \qquad i = 0, 1, 2 \qquad \ldots \quad (2.2)$$

where

$$\begin{aligned} \varphi_*^{(0)} &= \varphi && \text{if } 0 < \varphi < 1/(1+\tau_0) \\ &= 1/(1+\tau_0) && \text{otherwise} \\ \varphi_*^{(1)} &= \min[\varphi, 1/(1+\tau_0)] && \text{if } \varphi > 0 \\ &= 0 && \text{otherwise} \\ \varphi_*^{(2)} &= \min[|\varphi|, 1/(1+\tau_0)]. \end{aligned}$$

Note that $\hat{\mu}_*^{(0)}$ is the modification of $\hat{\mu}$ commonly used in practice and studied in Shah (1964, 1971). The three estimators $\hat{\mu}_*^{(0)}, \hat{\mu}_*^{(1)}, \hat{\mu}_*^{(2)}$ are identical if φ is non-negative a.s. as assumed in Shah (1971). In this case, $\hat{\mu}_*^{(0)}$ satisfies the condition of Theorem 2.3, part (i) of which, therefore, yields the result of Shah (1971). In the general situation e.g. if $\text{Prob}(-1/(1+\tau_0) < \varphi < 0) > 0$, the estimator $\hat{\mu}_{\cdot}^{(0)}$ does not satisfy the condition of Theorem 2.3. Hence the estimators $\hat{\mu}_{\cdot}^{(1)}, \hat{\mu}_{\cdot}^{(2)}$ which always satisfy the condition of Theorem 2.3, appear to be preferable to $\hat{\mu}_{\cdot}^{(0)}$. In the theorem which follows, we compare these three modifications of $\hat{\mu}$.

Theorem 2.4 : *Let* $V_i = V(\hat{\mu}_{\cdot}^{(i)})$, *where* $\hat{\mu}_{\cdot}^{(i)}$ *are as defined in* (2.1). *Then*

(i) $V_2 \leqslant V_1$ *for all* $\tau < 1+2\tau_0$

(ii) $V_0 \leqslant V_1$ *for all* $\tau < 1+2\tau_0$
$\geqslant V_1$ *for all* $\tau > 1+2\tau_0$

(iii) $V_0 \leqslant V_2$ *for all* $\tau \leqslant \tau_0$
$\geqslant V_2$ *for all* $\tau \geqslant 1+2\tau_0$.

Furthermore, the inequality between each pair of V_i*'s as stated above holds strictly unless the corresponding estimators are identical almost sure.*

Proof : Let $c = 1/(1+\tau_0)$. Then to prove (i), it suffices to show that for all $t \in (0, c]$ we have $h(t) < h(0) = 0$ if $\tau < 1+2\tau_0$. To prove (ii), it suffices to show that $h(c) < h(0) = 0$ if $\tau < 1+2\tau_0$ and $h(c) > h(0) = 0$ if $\tau > 1+2\tau_0$. All these hold by Lemma 2.1(c). To prove (iii) it suffices to show that for every $t \in (0, c)$, we have $h(t) > h(c)$ if $\tau \leqslant \tau_0$ and $h(t) > h(c)$ if $\tau \geqslant 1+2\tau_0$. This holds by Lemma 2.1(d).

Note : The referee has pointed out that the case where τ is small (say between τ_0 and $1+2\tau_0$) is more important than the case where τ is large since for large values of τ the maximum reduction in variance over $V(x)$ is small. Furthermore a negative value of ϕ would generally indicate that the true value of τ is rather small. In view of these the use of $\hat{\mu}_{\cdot}^{(0)}$ which (at the present state of our knowledge) appears to be hard to beat in the range of τ between τ_0 and $1+2\tau_0$, may be justified even in the general case when ϕ can assume negative values. On the other hand without denying these considerations fully, one may question if $1+2\tau_0$ can in all cases serve as the upper limit beyond which the maximum reduction in variance over $V(x)$ is negligible; and then, since a negative value of ϕ may be encountered with a positive probability even for large τ, one would, if possible, like to have a modified combined estimator which improves upon the original estimator for all $\tau \geqslant \tau_0$ and consequently retains the property of being uniformly better than the intra-block estimator for all $\tau \geqslant \tau_0$ whenever the original estimator has that property. Both $\hat{\mu}_{\cdot}^{(1)}$ and $\hat{\mu}_{\cdot}^{(2)}$ would serve this purpose but $\hat{\mu}_{\cdot}^{(0)}$ may not. Besides, $\hat{\mu}_{\cdot}^{(0)}$ is dominated by both $\hat{\mu}_{\cdot}^{(1)}$ and $\hat{\mu}_{\cdot}^{(2)}$ for all $\tau > 1+2\tau_0$ while the possibility of $\hat{\mu}_{\cdot}^{(2)}$ dominating $\hat{\mu}_{\cdot}^{(0)}$ for smaller values of τ remains open.

Acknowledgements. The author wishes to thank Professor S. K. Mitra for helpful discussion and valuable advice during the preparation of this paper.

References

Bose, R. C., Clatworthy, W. H. and Shrikhande, S. S. (1954): Tables of partially balanced designs with two associate classes. *North Carolina Agricultural Experimental Station Technical Bulletin* 107.

Khatri, C. G. and Shah, K. R. (1974): Estimation of location parameters from two linear models under normality. *Comm. Statist.*, **3**, 647-663.

Pearce, S. C., Calinski, T. and Marshall, T. Dec (1974): The basic contrasts of experimental design with special reference to the analysis of data. *Biometrika*, **64**, 449-460.

Shah, K. R. (1964): On a local property of combined inter and intra-block estimator. *Sankhyā*, A, **26**, 87-90.

——— (1971): Use of truncated estimator of variance ratio in recovery of inter-block information. *Ann. Math. Statist.*, **42**, 816-819.

Stein, C. (1966): An approach to the recovery of inter-block information in balanced incomplete block design. *Research Papers in Statistics*, ed. F. N. David. Wiley, New York, 351-366.

Yates, F. (1939): The recovery of inter-block information in varietal trials arranged in three dimensional lattices. *Ann. Eugenic*, **9**, 136-156.

Paper received : *June*, 1980.

Revised : *January*, 1983.