UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

A NOTE ON THE BEST LINEAR UNBIASED ESTIMATES FOR MULTIVARIATE POPULATIONS

by

J. N. Srivastava

November 1962

Contract No. AF 49(638)-213

In this note, we present a lemma on the best linear unbiased estimates for multivariate populations.

This research was supported by the Air Force Office of Scientific Research.

Institute of Statistics
Mimeo Series No. 339

A NOTE ON THE BEST LINEAR UNBIASED ESTIMATES FOR MULTIVARIATE POPULATIONS[1]

by

J. N. Srivastava
University of North Carolina

= =

Consider the usual multivariate linear model

(1) $$\underset{n\times p}{\mathrm{Exp}(Y)} = \underset{n\times m}{A}\ \underset{m\times p}{\xi} ,$$

where $n \geq m$ and where

(2) $$Y = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_p) = \begin{bmatrix} y_{11}, & y_{12}, & \ldots, & y_{1p} \\ \cdot & \cdot & \cdots & \cdot \\ y_{n1}, & y_{n2}, & \ldots, & y_{np} \end{bmatrix} = \begin{bmatrix} \underline{y}(1) \\ \vdots \\ \underline{y}(n) \end{bmatrix} , \text{ say}$$

is a matrix of np opservations; A is a known matrix and

(3) $$\xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1p} \\ \cdot & \cdot & \cdots & \cdot \\ \xi_{m1} & \xi_{m2} & \cdots & \xi_{mp} \end{bmatrix} = (\underline{\xi}_1, \underline{\xi}_2, \ldots, \underline{\xi}_p), \text{ say}$$

is a matrix of unknown parameters. We further assume that the vectors $\underline{y}_{(r)}$ (r = 1, 2, ... p) are all uncorrelated and that for r = 1, 2, ..., n:

(4) $$\mathrm{Var}\ (\underline{y}_{(r)}) = \underset{p\times p}{\Sigma} = (\sigma_{jj'}) , \text{ say} ,$$

where the dispersion matrix Σ is also unknown. The model (1) for the j-th variable reduces to

[1]This research was supported by the Air Force Office of Scientific Research.

(5) $$\text{Exp}\,(\underline{y}_j) = A\,\underline{\xi}_j$$

$$\text{Var}\,(y_{jr}) = \sigma_{jj}\,, \quad j = 1, 2, \ldots p, \quad r = 1, 2, \ldots n\,.$$

If we consider just the j-th variable and ignore the rest, we can obtain from (5), the best linear unbiased estimate $\underline{c}_j'\,\hat{\underline{\xi}}_j$ of $\underline{c}_j'\,\underline{\xi}_j$, where $\underline{c}_j$ is any mx1 vector such that $\underline{c}_j'\,\underline{\xi}_j$ is estimable. Let

(6) $$\Theta = \sum_{j=1}^{p} \underline{c}_j'\,\underline{\xi}_j$$

be a linear function of all the mp unknown parameters, such that for each j, $\underline{c}_j'\,\underline{\xi}_j$ is estimable. Let

(7) $$u = \sum_{j=1}^{p} \underline{c}_j'\,\hat{\underline{\xi}}_j\,.$$

Then we show that u is the best linear unbiased estimate of Θ.

<u>Lemma</u>: Let

$$z = \underline{b}_1'\,\underline{y}_1 + \ldots + \underline{b}_p'\,\underline{y}_p$$

be any other linear unbiased estimate of Θ. Then, provided that the space of the mpx1 vector

$$(\xi_{11}, \xi_{12}, \ldots, \xi_{1p}, \xi_{21}, \ldots\ldots, \xi_{mp})$$

contains at least mp linearly independent points, we must have

$$\text{Var}\,(z) > \text{Var}\,(u),$$

whatever the population dispersion matrix Σ may be. (Notice that no assumption of normality is involved.)

<u>Proof</u>: Suppose

(8) $$u = \underline{d}_1'\,\underline{y}_1 + \ldots\ldots + \underline{d}_p'\,\underline{y}_p\,.$$

Since

$$\text{Exp}\,(z) = E(u) = \Theta\,,$$

we have

$$\text{Exp}\ (z - u) = 0\ ,$$

or

$$(\underline{b}_1' - \underline{d}_1')\ A\ \underline{\xi}_1 + \dots + (\underline{b}_p' - \underline{d}_1')\ A\ \underline{\xi}_p \quad = \ 0\ ,$$

for all $\underline{\xi}_1, \underline{\xi}_2, \dots, \underline{\xi}_p$. This however implies

(9) $$(\underline{b}_j' - \underline{d}_j')\ A = 0_{1m}\ , \quad j = 1, 2, \dots, p\ ,$$

where 0_{1m} is a $1 \times m$ matrix. Also since $\underline{b}_j$ and $\underline{d}_j$ are free of the observations Y, we have

(10) $$\text{Var}\ (u) = \text{Var}\ (\underline{d}_1'\ \underline{y}_1 + \dots + \underline{d}_p'\ \underline{y}_p)$$

$$= \sum_{j=1}^{p} (\underline{d}_j'\ \underline{d}_j)\ \sigma_{jj} + \sum_{j \neq j'} (\underline{d}_{j'}'\ \underline{d}_j)\ \sigma_{jj'}\ ,$$

and similarly

$$\text{Var}\ (z) = \sum_{j=1}^{p} (\underline{b}_j'\ \underline{b}_j)\sigma_{jj} + \sum_{j \neq j'} (\underline{b}_{j'}',\ \underline{b}_j)\ \sigma_{jj'}\ .$$

Let A be of rank r and let $\overline{W}$ be the vector space of rank $n-r$, which is orthogonal to the columns of A. Let $\underline{\theta}_1, \underline{\theta}_2, \dots, \underline{\theta}_{n-r}$ be an orthogonal basis of $\overline{W}$. Then from (9), there exist constants $\mu_{j1}, \mu_{j2}, \dots, \mu_{j,n-r}$ $(j = 1, 2, \dots p)$ such that

(11) $$\underline{b}_j = \underline{d}_j + \mu_{j1}\ \underline{\theta}_1 + \mu_{j2}\ \underline{\theta}_2 + \dots + \mu_{j,n-r}\ \underline{\theta}_{n-r}, \quad j = 1, 2, \dots p\ .$$

Let W be the vector space of rank r (orthogonal to $\overline{W}$) generated by the columns of A. Then since $\underline{c}_j'\ \underline{\xi}_j$ is estimable as a univariate problem for the j-th variable, it follows that

$$\text{Rank}\ (A) = \text{Rank}\ \begin{pmatrix} A \\ \underline{c}_j' \end{pmatrix}\ , \quad j = 1, 2, \dots, p\ ,$$

and hence that $\underline{d}_j \in W$, for all j.

Hence we have from (11),

$$\underline{b}_j' \underline{b}_j = \underline{d}_j' \underline{d}_j + \mu_{j1}^2 + \mu_{j2}^2 + \cdots + \mu_{j,n-r}^2$$

$$\underline{b}_{j'}' \underline{b}_j = \underline{d}_{j'}' \underline{d}_j + \mu_{j1} \mu_{j1} + \cdots + \mu_{j,n-r} \mu_{j',n-r} .$$

Therefore we get

$$\text{Var } (z) - \text{Var } (u)$$

$$= \sum_{j=1}^{p} \left(\sum_{s=1}^{n-r} \mu_{js}^2 \right) \sigma_{jj} + \sum_{j \neq j'} \left(\sum_{s=1}^{n-r} \mu_{js} \mu_{j's} \right) \sigma_{jj'}$$

$$= \sum_{s=1}^{n-r} \left[\sum_{j=1}^{p} \mu_{js}^2 \sigma_{jj} + \sum_{j \neq j'} \mu_{js} \mu_{j's} \sigma_{jj'} \right]$$

$$= \sum_{s=1}^{n} \left[\underline{\mu}_s' \Sigma \underline{\mu}_s \right], \text{ where } \underline{\mu}_s' = (\mu_{1s}, \mu_{2s}, \ldots, \mu_{ps}) .$$

But since Σ is positive definite,

$$\underline{\mu}_s' \Sigma \underline{\mu}_s > 0 , \text{ unless } \underline{\mu}_s = 0_{1p} \text{ (zero vector)} .$$

Since however z is different from u, we must have $\underline{\mu}_s \neq 0_{1p}$, for some s. Hence

$$\text{Var } (z) > \text{Var } (u) ,$$

which proves the lemma.

The above lemma opens the door for many new lines of work in multivariate linear estimation. It has many interesting corrolaries, of which a very obvious one is the following:

Let $x_1, x_2, \ldots, x_n$ be independent random variables such that

$$\text{Exp } (X_j) = \theta_j ,$$

and $$\text{Var } (X_j) = \sigma_j^2 , \quad j = 1, 2, \ldots, n ,$$

where σ_j^2 are unknown and are not necessarily equal. Then if $a_1, a_2, \ldots, a_n$ are any real numbers, the best linear unbiased estimate of $\sum_{j=1}^{n} a_j \Theta_j$ is

$$\sum_{j=1}^{n} a_j \ y_j \ .$$

It is outside the scope of the present note to go into the details of applications of the above lemma. However, by way of illustration, two examples may be illuminating. A detailed paper will follow later.

Example 1. Consider a 2^n factorial experiment with r repetitions of each treatment combination. Also, assume no blocks to be present. Let $(g_1, g_2, \ldots, g_n)$, $g_j = 0$ or 1, $j = 1, 2, \ldots, n$, represent a treatment combination, $\Theta(g_1, g_2, \ldots, g_n)$ its 'true' effect and $y_i \ (g_1, g_2, \ldots, g_n)$ the corresponding i-th $(i = 1, 2, \ldots, r)$ observed value. Further suppose that all the $2^n r$ observations are independent, and that

$$\text{Var}\,[\,y_i(g_1, g_2, \ldots, g_n)\,] = \sigma^2(g_1, g_2, \ldots, g_n), \quad i = 1, 2, \ldots, r$$

depends on $g_1, g_2, \ldots, g_n$. Notice that this assumption is contrary to the usual one, where the variances are assumed to stay constant. However, the application of the lemma (in fact merely the corrollary) shows that if $i_1, i_2, \ldots, i_k$ are any k $(0 \leq k \leq n)$ factors, then the estimate of this k-factor interaction, say $A_{i_1} A_{i_2} \ldots A_{i_k}$ is the same contrast of $\bar{y} \ (g_1, g_2, \ldots, g_n)$, as $A_{i_1} A_{i_2} \ldots A_{i_k}$ is of $\Theta \ (g_1, g_2, \ldots, g_n)$. Here

$$\bar{y} \ (g_1, g_2, \ldots, g_n) = \frac{1}{r} \sum_{i=1}^{r} y_i \ (g_1, g_2, \ldots, g_n) \ .$$

In other words the best linear unbiased estimate of any interaction is unaltered by the assumption of unequal variances.

Example 2. As another area of application, consider an experiment, with say v treatments, repeated at different points of time. At any fixed point of time t, the observations are supposed to be independent, and have a variance σ_t^2

(dependent on t). Such a situation is obtained when for example, the treatments are the different rations for pigs, and we are studying their growth. Suppose τ and τ' are any two treatments, and the design is connected, and τ_t, τ'_t are their true effects. Then if we want to estimate something like

$$\theta = \sum_{t \in R} a_t (\tau_t - \tau'_t) \quad ,$$

where R is the total set of time points, the above lemma says that we could proceed by first obtaining the best linear unbiased estimate $(\hat{\tau}_t - \hat{\tau}'_t)$ for each $t \in R$, and then the best linear unbiased of θ is given by

$$\hat{\theta} = \sum_{t \in R} a_t (\tau_t - \tau'_t) \; .$$

I am thankful to Professor S. N. Roy for going through this note and for his comments.

REFERENCES

[1] Bose, R. C., "Notes on Linear Estimation", Unpublished Class Notes, University of North Carolina, Chapel Hill, N. C. (1958).

[2] Henry Scheffe: The Analysis of Variance, John Wiley and Sons, 1960.

[3] Roy, S. N., "Some Aspects of Multivariate Analysis," John Wiley and Sons, 1957.