A MULTIVARIATE EXTENSION OF THE GAUSS-MARKOV THEOREM[1)]

J. N. SRIVASTAVA

(Received June 9, 1964)

1. Summary

In this note we present the best linear unbiased estimates for multivariate populations, which may not necessarily be normal.

2. The multivariate extension

Consider the usual multivariate linear model

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

where $n \geqq m$ and where

$$Y = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_p) = \begin{bmatrix} y_{11}, & y_{12}, & \cdots, & y_{1p} \\ \cdot & \cdot & \cdots & \cdot \\ y_{n1}, & y_{n2}, & \cdots, & y_{np} \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} \boldsymbol{y}_{(1)} \\ \vdots \\ \boldsymbol{y}_{(n)} \end{bmatrix}, \text{ say}$$

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

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

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

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

where the dispersion matrix Σ is also unknown. The model (1) for the

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

jth variable reduces to

$$\text{(5)} \qquad \text{Exp}\,(\boldsymbol{y}_j) = A\boldsymbol{\xi}_j$$
$$\text{Var}\,(y_{jr}) = \sigma_{jj}\,, \qquad j=1, 2, \cdots, p\,, \quad r=1, 2, \cdots, n\,.$$

If we consider just the jth variable and ignore the rest, we can obtain from (5), the best linear unbiased estimate $\boldsymbol{c}_j'\hat{\boldsymbol{\xi}}_j$ of $\boldsymbol{c}_j'\boldsymbol{\xi}_j$, where $\boldsymbol{c}_j$ is any $m\times 1$ vector such that $\boldsymbol{c}_j'\boldsymbol{\xi}_j$ is estimable. Let

$$\text{(6)} \qquad \theta = \sum_{j=1}^{p} \boldsymbol{c}_j'\boldsymbol{\xi}_j$$

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

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

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

THEOREM. *Let*

$$z = \boldsymbol{b}_1'\boldsymbol{y}_1 + \cdots + \boldsymbol{b}_p'\boldsymbol{y}_p$$

be any other linear unbiased estimate of θ. *Then, provided that the space of the* $mp\times 1$ *vector*

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

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.)

PROOF. Suppose

$$\text{(8)} \qquad u = \boldsymbol{d}_1'\boldsymbol{y}_1 + \cdots + \boldsymbol{d}_p'\boldsymbol{y}_p\,.$$

Since

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

we have

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

or

$$(\boldsymbol{b}_1' - \boldsymbol{d}_1')A\boldsymbol{\xi}_1 + \cdots + (\boldsymbol{b}_p' - \boldsymbol{d}_p')A\boldsymbol{\xi}_p = 0\,,$$

for all $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_p$. This however implies

$$(9) \qquad (\boldsymbol{b}'_j - \boldsymbol{d}'_j)A = 0_{1m}, \qquad j=1, 2, \cdots, p,$$

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

$$(10) \qquad \begin{aligned} \mathrm{Var}\,(u) &= \mathrm{Var}\,(\boldsymbol{d}'_1\boldsymbol{y}_1 + \cdots + \boldsymbol{d}'_p\boldsymbol{y}_p) \\ &= n\sum_{j=1}^{p}(\boldsymbol{d}'_j\boldsymbol{d}_j)\sigma_{jj} + n\sum_{j\neq j'}(\boldsymbol{d}'_{j'}\boldsymbol{d}_j)\sigma_{jj'}, \end{aligned}$$

and similarly

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

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

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

Let W be the vector space of rank r (orthogonal to $\bar{W}$) generated by the columns of A. Then since $\boldsymbol{c}'_j\boldsymbol{\xi}_j$ is estimable as an univariate problem for the jth variable, it follows that

$$\mathrm{Rank}\,(A) = \mathrm{Rank}\begin{pmatrix} A \\ \boldsymbol{c}'_j \end{pmatrix}, \qquad j=1, 2, \cdots, p,$$

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

Hence we have from (11)

$$\begin{aligned} \boldsymbol{b}'_j\boldsymbol{b}_j &= \boldsymbol{d}'_j\boldsymbol{d}_j + \mu^2_{j1} + \mu^2_{j2} + \cdots + \mu^2_{j,\,n-r} \\ \boldsymbol{b}'_j\boldsymbol{b}_{j'} &= \boldsymbol{d}'_j\boldsymbol{d}_{j'} + \mu_{j1}\mu_{j'1} + \cdots + \mu_{j,\,n-r}\mu_{j',\,n-r}. \end{aligned}$$

Therefore we get

$$\begin{aligned} &\mathrm{Var}\,(z) - \mathrm{Var}\,(u) \\ &\quad = n\sum_{j=1}^{p}\Big(\sum_{s=1}^{n-r}\mu^2_{js}\Big)\sigma_{jj} + n\sum_{j\neq j'}\Big(\sum_{s=1}^{n-r}\mu_{js}\mu_{j's}\Big)\sigma_{jj'} \\ &\quad = n\sum_{s=1}^{n-r}\Big[\sum_{j=1}^{p}\mu^2_{js}\sigma_{jj} + \sum_{j\neq j'}\mu_{js}\mu_{j's}\sigma_{jj'}\Big] \\ &\quad = n\sum_{s=1}^{n}[\boldsymbol{\mu}'_s\Sigma\boldsymbol{\mu}_s], \text{ where } \boldsymbol{\mu}'_s = (\mu_{1s}, \mu_{2s}, \cdots, \mu_{ps}). \end{aligned}$$

But since Σ is positive definite,

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

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

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

which proves the theorem.

3. Acknowledgment

I am thankful to Professors S. N. Roy and Wassily Hoeffding for going through this note and for their comments.

UNIVERSITY OF NORTH CAROLINA[2)]

REFERENCES

[1] R. C. Bose, "Notes on linear estimation," *Unpublished Class Notes*, University of North Carolina, Chapel Hill, N.C., 1958.

[2] Henry Scheffé, *The Analysis of Variance*, John Wiley and Sons, 1960.

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

2) Now at the University of Nebraska.