AN ALTERNATIVE MULTIVARIATE PRODUCT ESTIMATOR

By

ASHOK SAHAI, RAM CHANDER AND ASHOK KUMAR MATHUR
Lucknow University, Lncknow.

(Received : September, 1978)

1. INTRODUCTION

Let Y_j ($j=1, 2, \ldots\ldots, N$) be the observation on jth population unit on the variable under study and X_{ij} ($j=1, 2, \ldots. N$; $i=1. 2, \ldots, p$) be the value of the ith auxiliary variable associated with jth population unit. Then, if a simple random sample of size n is drawn without replacement, let the corresponding sample means of these varlables be : $\bar{y}, \bar{x}_1, \ldots, \bar{x}_p$. Singh [5] studied the multivariable product estimator.

$$\bar{y}_{0p}=\sum_{i=1}^{p} w_i \frac{Z_i}{\bar{X}_i} \qquad \ldots(2.1)$$

where $Z_i=\bar{y}\bar{x}_i$; $\sum_{i=1}^{p} w_i=1$ and $w'_{(1\times p)} =(w_1, w_2, \ldots, w_p)$

He has obtained the optimum w_i's and found that upto order n^{-1},

$$E(\bar{y}_{0p})=\bar{Y}+\frac{1}{n} w'b \qquad \ldots(2.2)$$

where $b_{(p\times 1)}=\text{Col}\left(\frac{S_{01}}{\bar{X}_1}, \frac{S_{02}}{\bar{X}_2}, \ldots, \frac{S_{0p}}{\bar{X}_p}\right)=C^{-1} \underset{-}{S}_0$ where again

$$C_{(p\times p)}=\text{diag}\,(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p) \text{ and}$$

$$\underset{-}{S}'_{0(1\times p)}=S_{01}, S_{02}, \ldots, S_{0p})$$

$$S_{0i}=\frac{f}{N-1}\sum_{a=1}^{N} (Y_a-\bar{Y})\,(X_{ia}-\bar{X}_i)$$

$$S_{ij}=\frac{f}{N-1}\sum_{a=1}^{N}(X_{ia}-\bar{X}_i)(X_{ja}-\bar{X}_j)$$

where $f=\frac{N-n}{N}$ is the finite population correction.

With the help of (2.2) he found that bias and mse :

$$B(\bar{y}_{OP}) = \frac{1}{n}\, w'b \qquad \text{...(2.3)}$$

and

$$M(\bar{y}_{OP})=\frac{\bar{Y}^2}{n}\, w'Aw \qquad \text{...(2.3a)}$$

where

$$A = (a_{ij})\ ;\ \text{and}\ a_{ij} = (C_{oo}+C_{oi}+C_{oj}+C_{ij})$$

and

$$C_{00}=\frac{S^2y}{\bar{Y}^2} \quad C_{0j}=\frac{S_{0j}}{\bar{Y}\bar{X}_j},$$

$$C_{0i}=\frac{S_{0i}}{\bar{Y}\bar{X}_i} \text{ and } C_{ij}=\frac{S_{ij}}{\bar{X}_i\bar{X}_j}$$

and optimum $\underline{w}$ (say $\underline{w}_0$) is found to be

$$\underline{w}'_0=\underline{e}'\ A^{-1}/\underline{e}'\ A^{-1}\ \underline{e}$$

which gives the optimum $\bar{y}_{0_p}{}^*$

where

$$w_{0i}=\frac{\text{sum of elements in } i\text{th column of } A^{-1}}{\text{sum of all } p^2 \text{ elements of } A^{-1}}$$

Then

$$M(\bar{y}_{0p})=\frac{\bar{Y}^2}{n}\ \frac{1}{\underline{e}'A^{-1}\underline{e}} \qquad \text{...(2.4)}$$

We propose an alternative product estimator $\bar{Y}_{PG}$ on the pattern of Shukla's [4] multivariate ratio estimator

$$\bar{y}_{PG}=\bar{y}\ \frac{w_1\bar{x}_1+w_2\bar{x}_2+\ldots\ldots+w_p\bar{x}_p}{w_1\bar{X}_1+w_2\bar{X}_2+\ldots\ldots+w_p\bar{X}_p}$$

$$=\bar{y}\ \frac{\underline{w}'\underline{\bar{x}}}{\underline{w}'\underline{\bar{X}}}$$

where

$$\underline{\bar{X}}'_{(1\times p)}=(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p)$$

and

$$\underline{\bar{x}}'_{(1\times p)}=(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p$$

It has been shown that the MSE's of both Singh's [5] and our estimator are equal. Also absolute biases of these estimators are found to be equal for optimum choices of w_i's ; but the bias of the proposed estimator is shown to be less, whatever may be the value of w, than that of Singh's [5] estimator, otherwise.

2. Bias, Mse and Comparison

$$\bar{y}_{PG}=\bar{y}\ \frac{\underline{w}'\underline{\bar{x}}}{\underline{w}'\underline{\bar{X}}} \qquad \text{...(3.1)}$$

putting $\bar{y}=\bar{Y}+\epsilon_0$; $\bar{x}_i=\bar{X}_i+\epsilon_i$, where $E(\epsilon_0)=0$

and $E(\epsilon_1)=0 \;\forall\, i=1, 2, \ldots, p$

$$\bar{y}_{PG}=(\bar{Y}+\epsilon_0)\left(1+\frac{\underline{w}'\underline{\epsilon}}{\underline{w}'\underline{\bar{X}}}\right) \qquad \ldots(3.2)$$

where $\underline{\epsilon}'_{(1\times p)}=(\epsilon_1, \epsilon_2, \ldots, \epsilon_p)$

$$E(\underline{\epsilon})=0,\ E(\epsilon_0\underline{\epsilon})=\frac{\underline{S}_0}{n} \text{ and } E(\underline{\epsilon}\,\underline{\epsilon}')=\frac{S}{n}$$

where again $S_{(p\times p)}=(S_{ij})$

Taking expectatoin of (3·2), we have

$$E(\bar{y}_{PG})=\bar{Y}+\frac{1}{n}\ \frac{\underline{w}'\underline{S}_0}{\underline{w}'\underline{\bar{X}}} \qquad \ldots(3.3)$$

and

$$\text{Bias}(\bar{y}_{PG})=\frac{1}{n}\ \frac{\underline{w}'\underline{S}_0}{\underline{w}'\underline{\bar{X}}}$$

Also

$$M(\bar{y}_{PG})=E(\bar{y}_{PG}-\bar{Y})^2$$

$$=E\left(\epsilon_0+\bar{Y}\frac{\underline{w}'\underline{\epsilon}}{\underline{w}'\underline{\bar{X}}}+\frac{\underline{w}'\epsilon_0\underline{\epsilon}}{\underline{w}'\underline{\bar{X}}}\right)^2 \qquad \ldots(3.4)$$

Assuming $\left|\frac{\underline{w}'\underline{\epsilon}}{\underline{w}'\underline{\bar{X}}}\right|<1$ and neglecting the terms of order higher than (n^{-1}), equation (3·4) becomes

$$M(\bar{y}_{PG})=E\left(\epsilon_0^2+2\bar{Y}\frac{\underline{w}'\epsilon_0\underline{\epsilon}}{\underline{w}'\underline{\bar{X}}}+\bar{Y}^2\frac{\underline{w}'\underline{\epsilon}\,\underline{\epsilon}'\underline{w}}{\underline{w}'\underline{\bar{X}}\,\underline{\bar{X}}'\underline{w}}\right)$$

$$=\frac{1}{n}\left(S_y^2+2\bar{Y}\frac{\underline{w}'\underline{S}_0}{\underline{w}'\underline{\bar{X}}}+\bar{Y}^2\frac{\underline{w}'S\underline{w}}{\underline{w}'\underline{\bar{X}}\,\underline{\bar{X}}'\underline{w}}\right)$$

$$=\frac{\frac{Y^2}{n}\ \underline{w}'\left[\left(\frac{S_y^2}{\bar{Y}^2}\right)\underline{\bar{X}}\,\underline{\bar{X}}'+\left(\frac{\underline{S}_0}{\bar{Y}}\right)\underline{\bar{X}}'+\left(\frac{\underline{S}_0}{\bar{Y}}\right)\underline{\bar{X}}'+S\right]\underline{w}}{\underline{w}'\underline{\bar{X}}\,\underline{\bar{X}}'\underline{w}}$$

$$=\frac{\bar{Y}^2}{n}\ \frac{\underline{w}'B\underline{w}}{\underline{w}'\underline{\bar{X}}\,\underline{\bar{X}}'\underline{w}} \qquad \ldots(3.5)$$

where $B_{(p\times p)}=CAC$,

$C=\text{diag.}\,(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p)$

and $A=(a_{ij})$ defined earlier.

B is symmetric positive definite matrix. Now w_i's are chosen so as to minimise $M(\bar{y}_{PG})$; By Cauchy's Generalized inequality

$$\frac{x'Mx}{(x'y)(y'x)}\geqslant\frac{1}{y'M^{-1}y} \qquad \ldots(3.6)$$

where M is positive definite symmetry matrix. Equality holds iff $Mx=\theta Y$, where $(\theta \neq O)$ is scalar quantity. Putting $x=\underline{w}$, $M=B$ and $y=\bar{X}$, we have from (3·5) and 3·6).

$$M(\bar{y}_{PG})_{opt}=\frac{\bar{Y}^2}{n}\ \frac{1}{\underline{\bar{X}}'B'^{-1}\underline{\bar{X}}} \qquad \ldots(3.7)$$

and corresponding optimum value $\underline{w}_0=\theta B^{-1}\bar{X}$. Since θ is some non-zero scallar quantity ; whatever may be the value of θ it does not affect the MSE. We choose $\theta=1$ than $w=B^{-1}\underline{\bar{X}}$.

where $B=CAC$ and $A=(a_{ij})$ defined earlier.

Hence optimum $\underline{w}=C^{-1}A^{-1}, \underline{e}=\underline{w}^*$ (say)

where $\underline{e}'_{(1\times p)}=(1, 1, \ldots, 1)$.

with the optimum value of $\underline{w}$ $(\underline{w}^*)$ we have

$$B(\bar{y}_{PG})=\frac{1}{n}\ \frac{\underline{e}'A^{-1}C^{-1}\underline{S}_0}{\underline{e}'A^{-1}C^{-1}\underline{\bar{X}}}$$

$$=\frac{1}{n}\ \frac{\underline{e}'A^{-1}\underline{b}}{\underline{e}'A^{-1}C^{-1}\underline{\bar{X}}}$$

$$=\frac{1}{n}\ \frac{\underline{e}'A^{-1}\underline{b}}{\underline{e}'A^{-1}\underline{e}} \qquad \ldots(3.8)$$

and

$$M(\bar{y}_{PG})=\frac{\bar{Y}^2}{n}\ \frac{\underline{e}'A^{-1}\underline{e}}{(\underline{e}'A^{-1}\underline{e})(\underline{e}'A^{-1}\underline{e})}$$

$$=\frac{\bar{Y}^2}{n}\ \frac{1}{\underline{e}'A^{-1}\underline{e}} \qquad (3.9)$$

By comparing (3.9) and 2.4) we see that $\bar{y}_{OP}$ and $\bar{y}_{PG}$ have identical MSE's.

For comparing the squares of the two biases, *i.e.*, (3.8) and (2.3) for respective optimum value of $\underline{w}$ ($\underline{wo}$ and $\underline{w}^*$) ; we have

$$[\text{Biase } (\bar{y}_{OP})]=\frac{1}{n^2}\ \frac{\underline{e}'A^{-1}\underline{b}\underline{b}'A^{-1}\underline{e}}{(\underline{e}'A^{-1}\underline{e})^2} \qquad (3.10)$$

by comparing (3.10) with the square of equation (3.8) we see that both biases have the same square for optimum value of $\underline{w}$.

The optimum $\underline{w}$ ($\underline{wo}$ and $\underline{w}^*$) may be calculated. Only in terms of A^{-1} which is unknown. Thus for finding optimum $\underline{w}$ we have to estimate A^{-1} from sample using sample counter-parts of a_{ij}'s. In that case the absolute biases of $\bar{y}_{OP}$ and $\bar{y}_{PG}$ may be different. In what follows, we have shown that the absolute bais of the proposed estimators $\bar{y}_{PG}$ is always less than that of $\bar{y}_{OP}$. First we state and prove a result which is used in the comparison of the two biases.

Lemma : If X and Y are two negatively correlated random variables then

$$[E(X/Y)]^2 > \left[\frac{E(X)}{E(Y)}\right]^2 .$$

Proof : Since X and Y are two negatively correlated random variables then if we shall consider the random variables (X/Y) and Y; these two variables will be negatively correlated (as increase in Y means a decrease in (X/Y) that is,

$$COV\,(X/Y,\, Y) < 0 \text{ or } E\left(\frac{X}{Y}\,.\,Y\right) - E\left(\frac{X}{Y}\right).\,E\,(Y) < 0$$

i.e. $E\,(X) < E(Y)E\left(\frac{X}{Y}\right)$

i.e., $$\left[\frac{E(X)}{E(Y)}\right] < E\left(\frac{X}{Y}\right)$$

or $$\left[\frac{E(X)}{E(Y)}\right]^2 < \left[E\left(\frac{X}{Y}\right)\right]^2$$

i.e., $$\left[\frac{\text{Mean of } X}{\text{Mean of } Y}\right]^2 < [\text{Mean of } (X/Y)]^2$$

Hence the Lemma.

Now since we have,

$$B(\bar{y}_{OP}) = \frac{1}{n}\, w'b = \frac{\frac{1}{n}\sum_{i=1}^{p} w_i \frac{S_{0i}}{\bar{X}_i}}{\sum_{i=1}^{p} w_i}$$

$$= \frac{1}{n}[\text{Mean of } (S_{0i}/\bar{X}_i)\text{'s}] \qquad (3.11)$$

and

$$B(\bar{y}_{PG}) = \frac{1}{n}\,\frac{\sum_{i=1}^{p} w_i S_{0i}}{\sum_{i=1}^{p} w_i \bar{X}_i}$$

$$= \frac{1}{n}\left[\frac{\text{Mean of } S_{0i}\text{'s}}{\text{Mean of } \bar{X}_i\text{'s}}\right] \qquad (3.12)$$

It is intuitionally and vividly clear that an increase in $\bar{X}_i$'s will never accompany an increase in $\frac{So_i\text{'s}}{\bar{X}_i}$. Hence $\frac{So_i\text{'s}}{\bar{X}_i}$ and $\bar{X}_i$'s are negatively correlated. Hence by applying the preceding result, we have

$$\left[\frac{\text{Mean of } So_i\text{'s}}{\text{Mean of } \bar{X}_i\text{'s}}\right]^2 < [\text{Mean of } (SOi/\bar{X}_i)\text{'s}]^2$$

i.e..
$$\left[\frac{\sum_{i=1}^{p} w_i' so_i}{\sum_{i=1}^{p} w_i \bar{X}_i}\right]^2 < \left[\sum_{i=1}^{p} w_i \frac{So_i}{\bar{X}_i}\right]^2$$

which shows that the bias of the proposed estimator is less than that of Singh's [5] estimator.

3. Remarks

Analogous to the advantage of Shukla's (4) estimator over Olkin's (3) estimator, our estimator is simpler to calculate. Also it is less biased than Singh's (5) multivariate product estimator except when $\underline{wo}$ and $\underline{w^*}$ are known. But in practice $\underline{wo}$ and $\underline{w^*}$ are unknown and have to be estimated as the elements of $\underline{A}$ are unknown.

Summary

Whenever supplementary information is suitably used, the estimator using the supplementary variable is more efficient than that based on single (main) variable only. In practice, Statistician may have more than one auxiliary variable correlated with the characteristic under study. Olkin [3] initiated the ratio estimator based on multi-auxiliary variables and found it more efficient than that based on single auxiliary variable. An alternative to this estimator was given by Shukla [4). Singh [5] introduced the idea of multi-variate product estimator when the auxiliary variables are highly negatively correlated with the main variable under study. We have proposed and studied a simpler multivariate product estimator on Shukla's [4] pattern.

Acknowledgement

The author is thankful to the Referee whose comments have led to many improvements in the Paper.

REFERENCES

[1] Cochran, W.G. (1962) : *Sampling Techniques* John Wiley & Sons., New York.

[2] Murthy, M.N. (1964) : On product estimation *Sankhya A*, 26, 69-74.

[3] Olkin, I. (1958) : On multivariate ratio estimator for finite populaion *Biometrika* 45, 154-165.

[4] Shukla, G.K. (1966) : An alternative multivariate ratio estimate for finite population, *Calcutta Statistical Association Bulletin* Vol. 15, pp 127-134.

[5] Singh, M.P. (1967) : Multivariate product method of estimators for finite population. *Journal of Ind. Soc. of Agri. Stat.* Vol. XIX-No. 2.