A GENERALIZED RATIO—PRODUCT—DIFFERENCE (RPD) ESTIMATOR IN DOUBLE SAMPLING

By

S. K. Ray and R. Karan Singh

Lucknow University

(Received: August, 1977)

Introduction

Consider simple random sampling without replacement, in two phases, from a population of N units for which the main and auxiliary characters y and x have population means $\bar{Y}$ and $\bar{X}$, respectively. Let us suppose that $\bar{x}'$, based on a first phase sample of size n', is the mean estimator of $\bar{X}$, and $\bar{x}$, $\bar{y}$, based on a second phase sample of size n, are the mean estimators of $\bar{X}$, $\bar{Y}$ respectively. For estimating $\bar{Y}$, we propose the *RPD* estimator

$$\bar{y}_{RPD}=(\bar{x}'/\bar{x})^k \{\bar{y}+\beta\,(\bar{x}^a-\bar{x}'^a)\}$$

where k, β and a are some appropriate constants.

We observe that, in double sampling, $\bar{y}_{RPD}$ reduces to ratio estimator $\bar{y}_{Rd}=\bar{y}(\bar{x}'/\bar{x})$ for $k=1$, $a=0$,
product estimator $\bar{y}_{Pd}=\bar{y}(\bar{x}/\bar{x}')$ for $k=-1$, $a=0$,
difference estimator $\bar{y}_{Dd}=\bar{y}+\beta\,(\bar{x}-\bar{x}')$ for $k=0$, $a=1$

Bias and Mean Square Error

We evaluate the sampling bias and mean square error of $\bar{y}_{RPD}$ to the first degree of approximation (*i.e.* upto the terms of order $0(n^{-1})$). Let

$$e_o=(\bar{y}-\bar{Y})/\bar{Y},\; e_1=(\bar{x}-\bar{X})/\bar{X},\; e_1'=(\bar{x}'-\bar{X})/\bar{X}$$

so that, evidently

$$E(e_o)=o=E(e_1)=E(e_1'),\; E(e_o^2)=(f/n)C_Y^2,$$

$$E(e_1^2)=(f/n)C_X^2,\; E(e_1'^2)=(f'/n')\,C_X^2,\; E(e_oe_1)=(f/n)rC_YC_X,$$

$$E(e_oe_1')=(f'/n')rC_YC_X \text{ and } E(e_1e_1')=(f'/n')\,C_X^2$$

where

$$f=(N-n)/N,\ f'=(N-n')/N,\ C_Y(=S_Y/\bar{Y}) \text{ and } C_X\ (=S_X/\bar{X})$$

are the coefficients of variation of y, x, respectively and r is the coefficient of correlation between them. Also, define

$$C=rC_Y/C_X,\ R=\bar{Y}/\bar{X} \text{ and } B=rS_Y/S_X$$

It is assumed that the sample is large enough to make $|e_o|$, $|e_1|$ and $|e_1'|$ sufficiently small to justify the first degree of approximation wherein we ignore the terms involving e's in degree greater than two. We have

$$E(\bar{y}_{RPD})=E[\{(1+e_1')/(1+e_1)\}^k\ \{\bar{Y}(1+e_o)+\beta(\bar{X})^a((1+e_1)^a - (1-e_1')^a)\}]$$

which gives, on expansion in powers of e's and simplification, to the first degree of approxination

$$\text{Bias }(\bar{y}_{RPD})=(1/n-1/n')\ [\bar{Y}k\ \{C_X^2\ (k+1)/2-rC_YC_X\} + \beta(\bar{X})^a\ a\{(a-1)/2-k\}\ C_X^2\ ] \quad ..\ (2.1)$$

It is clear that, for suitable choice of a and k, the bias will be negligible for sufficiently large n.

Again, we have

$$MSE(\bar{y}_{RPD})=E(\bar{y}_{RPD}-\bar{Y})^2$$

which gives, on expension in powers of e's and simplification, to the first degree of approximation

$$MSE(\bar{y}_{RPD})=(1/n-1/n')\ \{(\bar{Y})^2\ (C_Y^2\ +k^2C_X^2\ -2krC_YC_X) + \beta^2a^2\ (\bar{X})^{2a}C_X^2\ +2\beta a\ \bar{Y}(\bar{X})^a\ (-kC_X^2\ +rC_YC_X)\} + (f'/n')\ (\bar{Y})^2C_Y^2 \quad ...(2.2)$$

From (2.2) we get at once

$$MSE(\bar{y}_{Rd})=(1/n-1/n')\ (\bar{Y})^2(C_Y^2\ +C_X^2\ -2rC_YC_X) + (f'/n')\ (\bar{Y})^2\ C_Y^2 \quad ...(2.3)$$

$$MSE(\bar{y}_{Pd})=(1/n-1/n')(\bar{Y})^2(C_Y^2\ +C_X^2\ +2rC_YC_X) + (f'/n')\ (\bar{Y})^2\ C_Y^2 \quad ...(2.4)$$

$$MSE\ (\bar{y}_{Dd}) = (1/n - 1/n')\ (\beta^2\ (\bar{X})^2\ C_X^2 + 2\beta\ (\bar{Y}\ \bar{X})\ rC_Y C_X) + (f/n)\ (\bar{Y})^2\ C_Y^2 \quad ...(2.5)$$

COMPARISON OF $\bar{y}_{RPD}$ WITH OTHER ESTIMATORS

We give below the conditions under which the mean square error of $\bar{y}_{RPD}$ is less than that of ratio, product of difference estimator in double sampling.

From (2.2) and (2.3) we have

$$MSE\ (\bar{y}_{RPD}) < MSE\ (\bar{y}_R)$$

if $$(k^2-1)\ R^2 - 2\ (k-1)\ RB + a^2\beta^2\ (\bar{X}^2)^{(a-1)} + 2a\beta\ (B-kR)\ (\bar{X})^{(a-1)} < 0 \quad ...(3.1)$$

From (2.2) and (2.4) we have

$$MSE\ (\bar{y}_{RPD}) < MSE\ (\bar{y}_{Pd})$$

if $$(k^2-1)\ R^2 - 2\ (k+1) RB + a^2\ \beta^2\ (\bar{X})^{2\ (a-1)} + 2a\beta\ (B-kR)\ (\bar{X})^{(a-1)} < 0 \quad ...(3.2)$$

And from (2.2) and (2.5) we have

$$MSE\ (\bar{y}_{RPD}) < MSE\ (\bar{y}_{Dd})$$

if $$\beta^2\ (a^2\ (\bar{X})^{2\ (a-1)} - 1) + 2\beta\ (a\ (B-kR)\ (\bar{X})^{(a-1)} - B) + k^2\ R^2 - 2kRB) < 0 \quad ...(3.3)$$

SUBCLASS OF RPD ESTIMATORS

We observe that the conditions (3.1), (3.2) and (3.3) may be simplified considerably by suitable choice of k, β and a. As an illustration, we take $k=1$, $\beta=1$ and $k=1$, $\beta=-1$ and consider the following two sub-classes of *RPD* estimators :

$$\bar{y}_* = (\bar{x}'/\bar{x})\ \{\bar{y} + (\bar{x}^a - \bar{x}'^a)\}$$

and

$$\bar{y}_{**} = (\bar{x}'/\bar{x})\ \{\bar{y} - (\bar{x}^a - \bar{x}'^a\}$$

Substituting $k=1$, $\beta=1$ in (3.1) we have

$$MSE\ (\bar{y}_*) < MSE\ (\bar{y}_{Rd})$$

if $$a^2\ (\bar{X})^{2\ (a-1)} + 2a\ (B-R)\ (\bar{X})^{(a-1)} < 0$$

or

if $$a\ (a + 2\ (B-R)\ (\bar{X})^{(1-a)}) < 0 \quad ...(4.1)$$

The condition (4.1) reduces to

$$rC_Y/C_X > 1 - a/(2R\ (\bar{X})^{(1-a)})\ \text{for}\ a<0 \quad ...(4.2)$$

and

$$rC_Y/C_X<1-a/\,(2R\,(\bar{X})^{(1-a)})\ \text{for}\ a>0 \qquad (4.3)$$

Again, substituting $k=1$, $\beta=-1$ in (3.1) we have

$$MSE\,(\bar{y}_{**}) < MSE\,(\bar{y}_{Rd})$$

if $\quad a^2\,(\bar{X})^{2\,(a-1)}-2a\,(B-R)\,(\bar{X})^{(a-1)}<0$

or

if $\quad a\,(a-2\,(B-R)\,(\bar{X})^{(1-a)})<0 \qquad (4.4)$

The condition (4.4) reduces to

$$rC_Y/C_X>1+a/(2R\,(\bar{X})^{(1-a)})\ \text{for}\ a>0 \qquad (4.5)$$

and

$$rC_Y/C_X<1+a/(2R\,(\bar{X})^{(1-a)})\ \text{for}\ a<0 \qquad (4.6)$$

CONCLUDING REMARKS

The efficiency conditions (4.2), (4.3), (4.5) and (4.6) contain the quantity $a/(2R\,(\bar{X})^{(1-a)})$ which can be made negligible for suitable choice of a. Due to the negligibility of $\frac{a}{2R\,\bar{X}^{(1-a)}}$, the conditions (4.2), (4.5) practically become $\frac{rC_X}{C_Y}>1$ and conditions (4.3), (4.6) practically become $rC_Y/C_X<1$. It is now clear that knowing whether $\frac{rC_Y}{C_X}=C$ is less or greater than unity we can choose $\bar{y}_*$ or $\bar{y}_{**}$ which will be better than $\bar{y}_{Rd}$. Analogously, we may also choose *RPD* estimators which are better than $\bar{y}_{Pd}$ and $\bar{y}_{Dd}$ in practical situations.

SUMMARY

A generalized estimator for estimating population mean, using observations on an auxiliary variable, is given for simple random sampling without replacement in two phases, of which the usual ratio, product and difference estimators in double sampling are special cases. The conditions for which the generalized estimator is better than the other estimators are obtained.

ACKNOWLEDGEMENT

The authors are grateful to the referees for making valuable comments for revising the paper.

REFERENCES

[1] Cochran, W.G. (1963) : Sampling Techniques, Second Edition, Wiley Eastern Private Limited, New Delhi.

[2] Sukhatme, P.V. and Sukhatme, B.V. (1976) : Sampling Theory of Surveye with Applications, Indian Society of Agricultural Statistics, New Delhi, India.