A NOTE ON GENERALIZED RPD ESTIMATOR IN DOUBLE SAMPLING

By

SURENDRA K. SRIVASTAVA

Punjabi University, Patiala

(Received : September, 1979)

SUMMARY

In this note it is shown that the generalized ratio-product-difference estimator in douule sampling proposed by Ray and Singh (1979) attains the same asymptotic minimum mean square error as that attained by a much simpler estimator defined by Srivastava (1970), and hence is unnecessarily complicated. Then an alternative equally simple estimator—a linear function of the si ple mean per unit estimator and ratio estimator in double sampling, which lso attains the same asymptotic minimum mean square error, has been shown to have the bias which is twice that of the Srivastava's estimator.

INTRODCUTION

For estimating the mean $\bar{Y}$ of a certain characteristic y of a finite population of size N, use is made of information on a suitable auxiliary characteristic $\bar{X}$ correlated with y. Ratio, product and difference methods of estimation are amongst the most widely used methods which utilise the known value $\bar{X}$ of the characteristic x. In case $\bar{X}$ is not known but inf-ormation on the characteristic x could be collected cheaply, the method of double sampling is adopted in which $\bar{X}$ in the estimator is replaced by its estimate $\bar{x}'$, the mean based on a preliminary simple random samle of size n' selected from the given population. Then a second phase sample of size n is drawn.

Ray and Singh (1979) have suggssted a double sampling estimator involving three unknown parameters

$$\bar{y}_{RPD}=\left(\frac{x'}{x}\right)^k\left\{\bar{y}+\beta\,(\bar{x}^a-\bar{x}'^a)\right\} \qquad \ldots(1)$$

and called it generalized ratio-product-difference estimator. They computed the bias and mean square error (MSE) of $\bar{y}_{RPD}$ up to the terms of order n^{-1}, under the double sampling scheme when the second-phase sample of size n is a subsample of the first-phase sample of size n'. In (1), $\bar{y}$ and $\bar{x}$ respectively denote the sample means of y and x based on the second-phase sample of size n. In this note we shall follow the notations of the paper by Ray and Singh (1979). For simplicity we shall assume that the finite population correction terms $f=(N-n)/N$ and $f'=(N-n')/N$ could be ignored.

DISCUSSION

No doubt, the ratio, product and difference estimator are special cases of the suggested estimator (1). For $\beta=0$ or $a=0$, the estimator (1) reduces to the double samplihg estimator of Shrivastava (1970), which itself reduces to ratio or product estimator according as $k=1$ or $k=-1$; and for $k=0$ and $a=1$, the estimator (1) reduces to the difference estimator. However, the authors have not discussed the optimum choice for the values of the parameter k. β and a in their estimator (1). The MSE of the estimator $\bar{y}_{RPD}$. up to the terms of order n^{-1}, as obtained by Ray and Singh (1979) is given by

$$\text{MSE }(\bar{y}_{RPD})=\left(\frac{1}{n}-\frac{1}{n'}\right)\Big\{\bar{y}^2(C_y^2+k^2C_x^2-2k\gamma C_yC_x)+\beta^2a^2\bar{X}^{2a}C_x^2+2\beta a\,\bar{y}\bar{X}^a(-kC_x^2+\gamma C_yC_x)\Big\}+\frac{1}{n'}\bar{y}^2C_y^2. \quad \ldots(2)$$

The optimum choice for the values of the three parameters k, β and a should be those which minimize the MSE in (2). Rearranging terms in (2), we have

$$\text{MSE }(\bar{y}_{RPD})=\frac{1}{n}\bar{Y}^2C_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)\Big\{(\beta a\bar{X}^a-k\bar{Y})^2C_x^2+2(\beta a\bar{X}^a-k\bar{Y})\bar{Y})\gamma C_yC_x\Big\}$$

$$=\frac{1}{n}\bar{Y}^2C_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)\Big\{\Big[(\beta a\bar{X}^a-k\bar{Y})C_x+\bar{Y}\gamma C_y\Big]^2-\bar{Y}^2\gamma^2C_y^2\Big\} \quad \ldots(3)$$

which is minimized for all choices of k, β and a satisfying

$$(\beta a\bar{X}^a-k\bar{Y})C_x+\bar{Y}\gamma C_y=0$$

or,
$$k-\frac{\beta a\bar{X}^{a-1}}{R}=\gamma\frac{C_y}{C_x} \quad \ldots(4)$$

where $R=\bar{Y}/\bar{X}$, and the minimum value of the MSE is given by

$$\text{Min MSE }(\bar{y}_{PRD})=\frac{1}{n}\bar{Y}^2C_y^2\ (1-\gamma^2)+\frac{1}{n'}\bar{Y}^2C_y^2\gamma^2. \qquad \ldots(5)$$

From (4) it is seen that the optimum values of k, β and a which minimize the MSE of $\bar{y}_{RPD}$' depend upon the population values $\gamma\frac{C_y}{C_x}$, R and $\bar{X}$, and a good guess of these values must be available to enable one to use the estimator $\bar{y}_{RPD}$. Also for any given set of these population values, a large number of values of k, β and a will satisfy (4) and any one set of these values will result in the same MSE of $\bar{y}_{RPD}$.

The simplest choice for the values of the parameters k, β and a satisfying (4) is $\beta a=0$ and $k=\gamma\frac{C_y}{C_x}$ which results in the estimator considered by Srivastava (1970), namely

$$\bar{y}_{d_s}=\bar{y}\left(\frac{\bar{x}'}{\bar{x}}\right)^k \qquad \ldots(6)$$

Thus the introduction of the parameters β and a unnecessarily complicates the estimator (1) without reducing its MSE below (5). The obtimum choice for k in the estimator (6) is $\gamma\frac{C_y}{C_x}$ and the optimum choice for k, β and a in the estimator (1) involves not only $\gamma\ \frac{C_y}{C_x}$ but also R and $\bar{X}$. Hence the estimator $\bar{y}_{d_s}$ defined at (6) is much simpler as compared to the estimator $\bar{y}_{RPD}$ defined at (1), but attains the same minimum MSE given by (5).

The two subclasses of $RP\$D$ estimators considered by Ray and Singh (1979), namely $\bar{y}_*$ when $k=1$, $\beta=1$ and $\bar{y}_{**}$ when $k=1$, $\beta=-1$ are also not simple. For example, the estimator $\bar{y}_*$ has minimum MSE equal to (5) when a is chosen to satisfy

$$a\bar{X}^{a-1}=R\left(1-\gamma\frac{C_y}{C_x}\right)$$

and this also involves unknown population values of $\gamma\ \frac{C_y}{C_x}$, R and $\bar{X}$

An interesting result follows from the above discussion. Up to the terms of order n^{-1}, the MSE of the estimator of form (1) cannot be reduced below (5), which is the MSE of the linear regression estimator in double sampling. And also that the estimators of form (1) are unnecessarily complicated and a much simpler estimator (6) attains the same minimum MSE as that attained by (1).

An Alternative Estimator

An alternative simple estimator[1] of the ratio type in double sampling is

$$\tilde{y}=(1-w)\bar{y}+w\bar{y}\left(\frac{\bar{x}'}{\bar{x}}\right) \qquad ...(7)$$

The estimator $\tilde{y}$ has bias of the order n^{-1} and its MSE up to the terms of order n^{-1}, is given by

$$\text{MSE }(\tilde{y})=\frac{1}{n}\bar{Y}^2C_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}^2(w^2C_x^2-2w\gamma C_yC_x)$$

This MSE is minimized for $w=\gamma\frac{C_y}{C_x}$ and the minimum MSE is the same as given by the expression (5).

The estimators $\bar{y}_{ds}$ and $\tilde{y}$ defined at (6) (7) are muchs impler as compared to the estimator (1) and for optimnm choice of the parameters involved, attain the same minimum MSE. Both these estimators involve only a single unknown parameter each, whose optimum value is given by $\gamma\ \frac{C_y}{C_x}$.

Between these two, one may choose the estimator whose absolute bias is smaller. Up to the terms of order n^{-1}, the bias of these estimators are

$$B(\bar{y}_{ds})=\left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}k\left(\frac{k+1}{2}-\gamma\frac{C_y}{C_x}\right)C_x^2$$

and

$$B(\tilde{y})=\left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}w\left(1-\gamma\frac{C_y}{C_x}\right)C_x^2$$

which for the optimum choices of k and w (both equal to $\gamma\frac{C_y}{C_x}$) reduce to

$$B(\bar{y}_{ds})=\frac{1}{2}\ B(\tilde{y})=\frac{1}{2}\left(\frac{1}{n}-\frac{1}{n'}\right)\bar{Y}\gamma\frac{C_y}{C_x}\left(1-\gamma\frac{C_y}{C_x}\right)C_x^2$$

Hence the bias of the estimator $\bar{y}_{ds}$ is always smaller than the bias of the estimator $\tilde{y}$.

---

[1]This estimator in the case of simple random sampling, has been studied by Chakrabarty (1968).

REFERENCES

Chakrabarty, R.P. (1968) : Contributions to the theory of ratiotype estimators. Ph. D. thesis, Texas *A* and *M* University.

Ray, S.K. and Singh, R.K. (1979) : A generalized ratio-product-difference (RPD) estimator in double sampling. *Jour. Ind. Soc. Agri. Stat.* 31, 15-88.

Srivastava, S.K. (1970) : A two-phase sampling estimator in sample surveys. *The Aust. Jour. Stat.* 12, 23-27.

Sukhatme, P.V. and Sukhatme, B. V. (1970) : *Sampling Theory of Surveys with Applications* Indian Society of Agricultural Statistics, New Delhi.