A CLASS OF COST EFFICIENT ESTIMATORS OF POPULATION RATIO OF MEANS IN PRESENCE OF NON RESPONSE

SHASHI BHUSHAN, R. KARAN SINGH , NAZIA NAQVI

Abstract: In the presence of non response, a class of estimators of finite ratio of two population means is proposed when the population mean of auxiliary variable is not known; its bias and MSE are found. Sub - Class of optimum estimators in the sense of having minimum MSE is found and enhancing the practical utility, a sub - classes of estimators depending on estimated optimum value based on sample observations is also investigated in the presence of non response. The expressions of sample size and inverse subsample fraction have been worked out by minimizing the cost for given MSE, and further the expressions of sample size and inverse subsample fraction have been found out by minimizing the MSE for fixed cost. A comparative study is also done.

Keywords: Class of estimators, auxiliary information, non response.

Introduction: Many sample surveys are beset with the problem of non response, that is, the information cannot be obtained from all the units in the survey. An estimator based on such partial and incomplete information is generally biased and the results may be grossly misleading when the respondents differ from the non-respondents. In their seminal paper, Hansen and Hurwitz (1946) considered a technique of sub-sampling the non-respondents in order to adjust for the non-response bias in a mail survey. When the auxiliary information on a variable x is known in the form of its population mean $\bar{X}$, and the non-response is present, then the problem of estimation of population mean $\bar{Y}$ of the study variate y has been dealt by various authors including Rao (1986, 1987) and Khare and Srivastava (1993, 1997).

Assuming the population mean $\bar{X}$ of the auxiliary variate x to be unknown and in presence of non-response, Okafor and Lee (2000) and Tabasum and Khan (2004) have proposed to use double sampling (or two phase sampling) procedure to estimate the population mean $\bar{X}$ on the basis of a large first phase sample of size n' drawn from the finite population of size N by simple random sampling without replacement (SRSWOR). Then a second phase sub - sample of size n $(n<n')$ is drawn from n' by SRSWOR and the information on study variable y under investigation is measured on it.

Singh and Bhushan (2013) proposed a more cost efficient method instead of resorting to double sampling, if the population mean $\bar{X}$ of the auxiliary variable x is not known then we propose to estimate the population mean $\bar{X}$ on the basis of sample of size n drawn from the finite population of size N by simple random sampling without replacement (SRSWOR) assuming that the complete information on the auxiliary variable x is available to us. Let n_1 units provide information on study variables y_i and $i=1,2$; and n_2 units do not respond. Then utilizing Hansen and Hurwitz technique we select a sub sample from n_2 non-responding units of size $r=\frac{n_2}{k}, k>1$.

In this paper, we have proposed the ratio - product - difference (RPD) class of estimators for estimating the ratio of the two population means in the presence of non response

$$\hat{R}_a^* = \hat{R}^* \left\{1+a\left(\frac{\bar{x}^*-\bar{x}}{\bar{x}}\right)\right\}$$

$$=(1-a)\hat{R}^* + a\hat{R}^*\frac{\bar{x}^*}{\bar{x}} \qquad (1.1)$$

where a is the characterizing scalar to be determined suitably. $\bar{x}$ is the sample mean of x based on n units ; $\bar{y}_i^* = \frac{n_1}{n}\bar{y}_{i(1)} + \frac{n_2}{n}\bar{y}_{i(2)}'$ for $i=1,2$ and $\bar{x}^* = \frac{n_1}{n}\bar{x}_{(1)} + \frac{n_2}{n}\bar{x}_{(2)}'$ where $(\bar{y}_{i(1)}, \bar{x}_{(1)})$ and $(\bar{y}_{i(2)}, \bar{x}_{(2)})$ are the sample means based on n_1 units and the subsample means based on r units of the variables (y_i, x) for $i=1,2$ respectively.

2. Bias and Mean Square Error of the proposed estimator:

Let us consider

$$e_0 = \frac{(\bar{x}-\bar{X})}{\bar{X}} \;;\; e_0^* = \frac{(\bar{x}^*-\bar{X})}{\bar{X}} \quad e_1^* = \frac{(\bar{y}_1-\bar{Y}_1)}{\bar{Y}_1}$$

; $e_2^* = \frac{\left(\bar{y}_2^* - \bar{Y}_2\right)}{\bar{Y}_2}$

with $E(e_0) = E(e_0^*) = E(e_1^*) = E(e_2^*) = 0$ and

$$E(e_0^2) = \frac{(N-n)}{Nn} C_x^2$$

$$E(e_0 e_1^*) = \frac{(N-n)}{Nn} \rho_{01} C_x C_{y_1}$$

$$E(e_0^{*2}) = \frac{(N-n)}{Nn} C_x^2 + \frac{(k-1)N_2}{Nn} C_{x(2)}^{'2}$$

$$E(e_0 e_2^*) = \frac{(N-n)}{Nn} \rho_{02} C_x C_{y_2}$$

$$E(e_0^{*2}) = \frac{(N-n)}{Nn} C_x^2 + \frac{(k-1)N_2}{Nn} C_{x(2)}^{'2}$$

$$E(e_0^* e_1^*) = \frac{(N-n)}{Nn} \rho_{01} C_x C_{y_1} + \frac{(k-1)N_2}{Nn} \rho_{01}^{'} C_{x(2)}^{'} C_{y_{1(2)}}^{'}$$

$$E(e_2^{*2}) = \frac{(N-n)}{Nn} C_{y_2}^2 \frac{((k-1)N_2)}{Nn} C_{y_{2(2)}}^{'2}$$

$$E(e_0^* e_2^*) = \frac{(N-n)}{Nn} \rho_{02} C_x C_{y_1} + \frac{(k-1)N_2}{Nn} \rho_{02}^{'} C_{x(2)}^{'} C_{y_{2(2)}}^{'}$$

$$E(e_0 e_0^*) = \frac{(N-n)}{Nn} C_x^2$$

$$E(e_1^* e_2^*) = \frac{(N-n)}{Nn} \rho_{12} C_{y_1} C_{y_2} + \frac{((k-1)N_2)}{Nn} \rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}$$

Now,

$$\hat{R}_a^* = \hat{R}^* \left\{1 + \frac{a(\bar{x}^* - \bar{x})}{\bar{x}}\right\}$$

$$= (1-a)\hat{R}^* + a\hat{R}^* \frac{\bar{x}^*}{\bar{x}}$$

$$\hat{R}_a^* - R = R\{\left(e_1^* - e_2^* + e_2^{*2} - e_2^{*3} - e_1^* e_2^* + \ldots\right) + a\left(e_0^* - e_0 + e_0^2 - e_0 e_0^* + e_0 e_2^* - e_0^* e_2^* - e_0 e_1^* + \ldots\right)\} \quad (2.1)$$

Taking expectation on both sides of (2.1) and ignoring the error terms of degree greater than two, we have bias of $\hat{R}_a^*$ to the first degree of approximation, that is, upto the terms of order $o\left(\frac{1}{n}\right)$ is to be

$$Bias(\hat{R}_a^*) = R\left[\frac{(N-n)}{Nn}\{C_{y_2}^2 - \rho_{12} C_{y_1} C_{y_2}\} + \frac{(k-1)N_2}{Nn}\left\{C_{y_{2(2)}}^{'2} - \rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}\right\}\right] + Ra\left[\frac{(k-1)N_2}{Nn}\left\{\rho_{10}^{'} C_{y_{1(2)}}^{'} C_{x(2)}^{'} - \rho_{20}^{'} C_{y_{2(2)}}^{'} C_{x(2)}^{'}\right\}\right]$$

$$= Bias(\hat{R}) + \frac{(k-1)N_2}{Nn} R\left[\left\{C_{y_{2(2)}}^{'2} - \rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}\right\} + a\left\{\rho_{10}^{'} C_{y_{1(2)}}^{'} C_{x(2)}^{'} - \rho_{20}^{'} C_{y_{2(2)}}^{'} C_{x(2)}^{'}\right\}\right] \quad (2.2)$$

where $\hat{R} = \frac{\bar{y}_1}{\bar{y}_2}$ is the usual estimator of R.

Squaring both the sides of (2.1), ignoring the error terms of degree greater than two and taking expectation, we have the mean square error of $\hat{R}_a^*$ to the terms of order $o\left(\frac{1}{n}\right)$ to be

$$MSE(\hat{R}_a^*) = MSE(\hat{R}) + \frac{(k-1)N_2}{Nn} R^2 \left[\left\{C_{y_{1(2)}}^{'2} + C_{y_{2(2)}}^{'2} - 2\rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}\right\} + a^2 C_{x(2)}^{'2} + 2a\left\{\rho_{10}^{'} C_{y_{1(2)}}^{'} C_{x(2)}^{'} - \rho_{20}^{'} C_{y_{2(2)}}^{'} C_{x(2)}^{'}\right\}\right] \quad (2.3)$$

$$= MSE(\hat{R}) + \frac{(k-1)N_2}{Nn} R^2 \left[D_1 + a^2 C_{x(2)}^{'2} + 2aD_2 C_{x(2)}^{'}\right] \quad (2.4)$$

where $D_1 = \left\{C_{y_{1(2)}}^{'2} + C_{y_{2(2)}}^{'2} - 2\rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}\right\}$

and $D_2 = \left\{\rho_{10}^{'} C_{y_{1(2)}}^{'} - \rho_{20}^{'} C_{y_{2(2)}}^{'}\right\}$

From (2.4), the optimum value of a minimizing the mean square error of $\hat{R}_a^*$ is

$$a_{opt} = -\frac{D_2}{C_{x(2)}^{'}} \quad (2.5)$$

and the minimum mean square error is given by

$$MSE(\hat{R}_a^*)_{opt} = MSE(\hat{R}) - \frac{(k-1)N_2}{Nn}$$

$$R^2\left[D_2^{\ 2}-D_1\right] \qquad (2.6)$$

showing that, for optimum value of a, the estimator is more efficient than the usual ratio estimator $\hat{\bar{R}}$. The value $a_{opt}=-\frac{D_2}{C'_{x_{(2)}}}$ involves unknown parameters; hence, the estimator based on the optimum value of a lacks its practical utility. Therefore the alternative is to replace the optimum value

$$a_{opt}=-\frac{D_2}{C'_{x_{(2)}}}$$

$$=-\frac{\left\{\rho'_{10}C'_{y_{1(2)}}-\rho'_{20}C'_{y_{2(2)}}\right\}}{C'_{x_{(2)}}}$$

$$=-\frac{\bar{X}}{S'^2_{x_{(2)}}}\left\{\frac{S'_{y_{1(2)}x_{(2)}}}{\bar{Y}_1}-\frac{S'_{y_{2(2)}x_{(2)}}}{\bar{Y}_2}\right\} \qquad (2.7)$$

by estimated optimum c (based on sample observations) of a_{opt} given by

$$c=-\frac{\bar{x}^*}{s'^2_{x_{(2)}}}\left\{\frac{s'_{y_{1(2)}x_{(2)}}}{\bar{y}_1^*}-\frac{s'_{y_{2(2)}x_{(2)}}}{\bar{y}_2^*}\right\} \qquad (2.8)$$

Replacing a_{opt} of a by the estimated optimum

$$c=-\frac{\bar{x}^*}{s'^2_{x_{(2)}}}\left\{\frac{s'_{y_{1(2)}x_{(2)}}}{\bar{y}_1^*}-\frac{s'_{y_{2(2)}x_{(2)}}}{\bar{y}_2^*}\right\}$$

in $\hat{R}_a^*$, we get the estimator

$$\hat{R}_a^*=\hat{R}^*\left\{1+\frac{c(\bar{x}^*-\bar{x})}{\bar{x}}\right\} \qquad (2.9)$$

whose bias and mean square error are found below in the next section 3.

3. Bias and Mean Square Error of the estimator $\hat{R}_c^*$

Along with e_0, e_0^*, e_1^* and e_2^* already defined in section 2, let

$$e_3^*=\frac{\left(s'_{y_{1(2)}x_{(2)}}-S'_{y_{1(2)}x_{(2)}}\right)}{S'_{y_{1(2)}x_{(2)}}} \quad e_4^*=\frac{\left(s'_{y_{2(2)}x_{(2)}}-S'_{y_{2(2)}x_{(2)}}\right)}{S'_{y_{2(2)}x_{(2)}}}$$

$$e_5^*=\frac{s'^2_{x_{(2)}}-S'^2_{x_{(2)}}}{S'^2_{x_{(2)}}}$$

Now, we have

$$\hat{R}_a^*=\hat{R}^*\left\{1+\frac{c(\bar{x}^*-\bar{x})}{\bar{x}}\right\}$$

$$=(1-c)\hat{R}^*+c\hat{R}^*\frac{\bar{x}^*}{\bar{x}}$$

$$\hat{R}_c^*-R=R\Bigg[\left(e_1^*-e_2^*+e_2^{*2}-e_2^{*3}-e_1^*e_2^*+\ldots\right)+\frac{\left(e_0-e_0^*+e_0e_0^*-e_0e_2^*+\ldots\right)\bar{X}S'_{y_{1(2)}x_{(2)}}}{\bar{Y}_1S'^2_{x_{(2)}}}-\frac{\left(e_0-e_0^*+e_0e_0^*+e_0e_1^*-\ldots\right)\bar{X}S'_{y_{2(2)}x_{(2)}}}{\bar{Y}_2S'^2_{x_{(2)}}}\Bigg] \qquad (3.1)$$

Taking expectation on both sides of (3.1), ignoring the error terms of degree greater than two, we can easily see that the bias of $\hat{R}_c^*$ to the first degree of approximation is of order $o\left(\frac{1}{n}\right)$; hence, for sufficiently large value of n, the bias becomes negligible.

Further, squaring on both sides of (3.1), taking expectation and ignoring the error terms of degree greater than two, the mean square error of $\hat{R}_c^*$ to the first degree of approximation, that is, up to terms of order $o\left(\frac{1}{n}\right)$ is

$$MSE\left(\hat{R}_c^*\right)=MSE\left(\hat{R}\right)+\frac{(k-1)N_2}{Nn}R^2\left[D_1+D_2^2-2D_2^2\right]$$

$$=MSE\left(\hat{R}\right)-\frac{(k-1)N_2}{Nn}R^2\left[D_2^2-D_1\right] \qquad (3.2)$$

4. Determination of Optimum Values of n & k for fixed Precision

Let us consider a cost function for $\hat{R}_a^*$ as,

$$C=c\,n+c_1n_1+c_2r$$

where c is the cost per unit of the first attempt with the sample, n ; c_1 is the cost per unit for processing the respondent data at the first attempt in n_1 and c_2 is the cost per unit associated with the subsample, r of n_2.

Since the values of n_1 and r is not known until the first attempt is made, so the expected cost will be used in planning the survey. The expected values of n_1 and r are $W_1 n$ and $\frac{W_2 n}{k}$. Thus the expected cost is given by,

$$E(C)=C^*=n\left[c+c_1W_1+\frac{c_2W_2}{k}\right] \quad (4.1)$$

To determine the optimum values of n and k that minimize the cost for a fixed variance V_0, we consider the function,

$$\phi=C^*+\lambda\left\{MSE\left(\hat{R}_a^*\right)-V_0\right\}$$

$$\phi=n\left[c+c_1W_1+\frac{c_2W_2}{k}\right]$$
$$+\lambda\left[R^2\left\{\left(\frac{1}{n}-\frac{1}{N}\right)D_0+\frac{(k-1)N_2}{Nn}\right.\right.$$
$$\left.\left.\left[D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right]\right\}-V_0\right] \quad (4.2)$$

where λ is lagrange's multiplier.
Now differentiating (4.2) with respect to n and k, and on equating them with zero we get,

$$n=R\sqrt{\frac{\lambda\left[D_0+(k-1)W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}{c+c_1W_1+\frac{c_2W_2}{k}}} \quad (4.3)$$

$$\frac{n}{k}=R\sqrt{\frac{\lambda\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}}{c_2}} \quad (4.4)$$

On putting (4.3) in (4.4) we get,

$$k_{opt}=\sqrt{\frac{c_2\left[D_0-W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}{\left((c+c_1W_1)\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right)}} \quad (4.5)$$

Further substituting the values of n and k in the expression of MSE we get,

$$\sqrt{\lambda}=\frac{\sqrt{\left[D_0+(k-1)W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}}{\left[\frac{V_0}{R}+\frac{RD_0}{N}\right]}$$
$$\times\sqrt{\left(c+c_1W_1+\frac{c_2W_2}{k}\right)} \quad (4.6)$$

On using this value of λ we get the optimum value of n as,

$$n_{opt}=\frac{\left[D_0+(k-1)W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}{\left[\frac{V_0}{R^2}+\frac{D_0}{N}\right]} \quad (4.7)$$

On substituting the optimum value of n and k in (4.1) we get the minimum cost for fixed variance V_0 as,

$$C^*=\frac{\left[D_0+(k_{opt}-1)W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}{\left[\frac{V_0}{R^2}+\frac{D_0}{N}\right]}$$
$$\times\left(c+c_1W_1+\frac{c_2W_2}{k_{opt}}\right) \quad (4.8)$$

Ignoring the terms of order $\frac{1}{N}$ we get the expected minimum cost for fixed precision as,

$$C^*=\frac{R^2\left[D_0+(k_{opt}-1)W_2\left\{D_1+a^2C_{x_{(2)}}^{'2}+2aD_2C_{x_{(2)}}^{'}\right\}\right]}{V_0}$$
$$\times\left(c+c_1W_1+\frac{c_2W_2}{k_{opt}}\right) \quad (4.9)$$

5. Determination of Optimum Values of& k for fixed Cost

Let C_0 be the total cost (fixed) of the survey apart from overhead cost. The expected total cost of the survey apart from the overhead cost is given by

$$C^*=n\left[c+c_1W_1+\frac{c_2W_2}{k}\right] \quad (5.1)$$

where c is the cost per unit of the first attempt with the sample, n ; c_1 is the cost per unit for processing the respondent data at the first attempt in n_1 and c_2 is the cost per unit associated with the sub sample r of n_2.

The MSE of $\hat{R}_a^*$ can be expressed as,

$$MSE(\hat{R}_a^*)=R^2\left\{\left(\frac{1}{n}-\frac{1}{N}\right)D_0+\frac{(k-1)N_2}{Nn}\right.$$

$$\left[D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right]\Big\}$$

$$= R^2\left[\frac{D_0 - W_2\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}}{n} + \frac{kW_2\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}}{n} - \frac{D_0}{N}\right] \quad (5.2)$$

To determine the optimum values of n and k that minimize the $MSE(\hat{R}_a^*)$ for a fixed cost $(C^* < C_0)$, we consider the function,

$$\phi^* = MSE(\hat{R}_a^*) + \lambda\left\{C^* - C_0\right\}$$

$$\phi^* = R^2\left\{\left(\frac{1}{n} - \frac{1}{N}\right)D_0 + \frac{(k-1)N_2}{Nn}\left[D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right]\right\} + \lambda\left\{n\left[c + c_1W_1 + \frac{c_2W_2}{k}\right] - C_0\right\} \quad (5.3)$$

where λ is lagrange's multiplier.
Now differentiating (5.3) with respect to n and k; and on equating them with zero we get

$$n = R\sqrt{\frac{\left[D_0 + (k-1)W_2\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\right]}{\lambda\left\{c + c_1W_1 + \frac{c_2W_2}{k}\right\}}} \quad (5.4)$$

$$\frac{n}{k} = R\sqrt{\frac{\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}}{\lambda c_2}} \quad (5.5)$$

On using (5.4) in (5.5) we get,

$$k_{opt} = \sqrt{\frac{c_2\left[D_0 - W_2\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\right]}{\left((c + c_1W_1)\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\right)}} \quad (5.6)$$

Further substituting the values of n and k in the expression of expected cost, we get,

$$\sqrt{\lambda} = R\sqrt{\frac{\left[D_0 + (k_{opt} - 1)W_2\left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\right]}{C^*}} \quad (5.7)$$

Further on substituting the value of λ in (5.4), we get the optimum value of n as,

$$n_{opt} = \frac{C^*}{\left\{c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right\}} \quad (5.8)$$

On substituting the optimum value of n and k, we get the mean square error of $\hat{R}_a^*$ for fixed cost $(C^* < C_0)$ as,

$$MSE\left(\hat{R}_a^*\right) = R^2\left\{\frac{\left[D_0 + (k_{opt} - 1)W_2\right.}{C^*} \times \left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\Big] \times \left(c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right) - \frac{D_0}{N}\right\} \quad (5.9)$$

Ignoring the terms of order $\frac{1}{N}$ we get the mean square error of $\hat{R}_a^*$ for fixed cost $(C^* < C_0)$ as,

$$MSE\left(\hat{R}_a^*\right) = R^2\left\{\frac{\left[D_0 + (k_{opt} - 1)W_2\right.}{C^*} \times \left\{D_1 + a^2C_{x_{(2)}}^{'2} + 2aD_2C_{x_{(2)}}^{'}\right\}\Big] \times \left(c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right)\right\} \quad (5.10)$$

6. A Comparative Study

It may be noted that the ratio estimator $\hat{R}^* = \frac{\bar{y}_1^*}{\bar{y}_2^*}$ and the ratio-cum-product estimator $\hat{R}_1^* = \hat{R}^*\frac{\bar{x}^*}{\bar{x}}$ are special cases of the proposed class of estimator for $a = 0$ and $C^* < C_{\hat{R}}^*$ respectively.

1. On comparing the proposed class of estimator $\hat{R}_a^*$ with the usual ratio estimator $\hat{R}^*$ and the ratio-cum-product estimator $\hat{R}_1^*$, it can be easily verified that,

$$MSE\left(\hat{R}^*\right) = MSE\left(\hat{R}\right) + \frac{(k-1)N_2}{Nn}R^2D_1 \quad (6.1)$$

$$MSE\left(\hat{R}_1^*\right) = MSE\left(\hat{R}\right) + \frac{(k-1)N_2}{Nn} R^2\left[D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right]$$

(6.2)

where $D_1 = \left\{C_{y_{1(2)}}^{'2} + C_{y_{2(2)}}^{'2} - 2\rho_{12}^{'} C_{y_{1(2)}}^{'} C_{y_{2(2)}}^{'}\right\}$ and

$D_2 = \left\{\rho_{10}^{'} C_{y_{1(2)}}^{'} - \rho_{20}^{'} C_{y_{2(2)}}^{'}\right\}$

there by showing that the proposed class of estimator is always more efficient than the ratio estimator $\hat{R}^*$ and the ratio-cum-product estimator $\hat{R}_1^*$

2. Further on comparing the proposed class of estimator $\hat{R}_a^*$ with the usual ratio estimator $\hat{R}^*$ and the ratio-cum-product estimator $\hat{R}_1^*$, it is seen that the expected minimum cost for fixed precision V_0 comes out to be,

$$C^* = \frac{R^2\left[D_0 + \left(k_{opt} - 1\right)W_2 D_1\right]\left(c + c_1 W_1 + \frac{c_2 W_2}{k_{opt}}\right)}{V_0} \quad (6.3)$$

where $k_{opt} = \sqrt{\frac{c_2\left[D_0 - W_2 D_1\right]}{\left\{\left(c + c_1 W_1\right) D_1\right\}}}$

$$C^* = \frac{R^2\left[D_0 + \left(k_{opt} - 1\right)W_2\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right]}{V_0} \times \left(c + c_1 W_1 + \frac{c_2 W_2}{k_{opt}}\right) \quad (6.4)$$

where $k_{opt} = \sqrt{\frac{c_2\left[D_0 - W_2\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right]}{\left(\left(c + c_1 W_1\right)\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right)}}$

It can be easily verified that for $a = a_{opt}$,

$C^* < C_{\hat{R}^*}^*$ and $C^* < C_{\hat{R}_1^*}^*$

3. Similarly on comparing the proposed class of estimator $\hat{R}_a^*$ with the usual ratio estimator $\hat{R}^*$ and the ratio-cum-product estimator $\hat{R}_1^*$, it is seen that the mean square error for fixed cost C_0 comes out to be,

$$MSE\left(\hat{R}^*\right) = \frac{R^2\left[D_0 + \left(k_{opt} - 1\right)W_2 D_1\right]}{C^*} \times \left(c + c_1 W_1 + \frac{c_2 W_2}{k_{opt}}\right) \quad (6.5)$$

where $k_{opt} = \sqrt{\frac{c_2\left[D_0 - W_2 D_1\right]}{\left\{\left(c + c_1 W_1\right) D_1\right\}}}$

$$MSE\left(\hat{R}_1^*\right) = \frac{R^2\left[D_0 + \left(k_{opt} - 1\right)W_2\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right]}{C^*} \times \left(c + c_1 W_1 + \frac{c_2 W_2}{k_{opt}}\right) \quad (6.6)$$

where

$$k_{opt} = \sqrt{\frac{c_2\left[D_0 - W_2\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right]}{\left(\left(c + c_1 W_1\right)\left\{D_1 + C_{x_{(2)}}^{'2} + 2D_2 C_{x_{(2)}}^{'}\right\}\right)}}$$

It can be easily verified that for $a = a_{opt}$,

$MSE\left(\hat{R}_a^*\right) < MSE\left(\hat{R}^*\right)$ and

$MSE\left(\hat{R}_a^*\right) < MSE\left(\hat{R}^*\right)$

Therefore, if the population mean $\bar{X}$ is not known, then the proposed class of estimator in presence of non-response is more effective as in this case we do not have to resort to double sampling.

7. An Empirical Study

The present data belong to the data on physical growth of upper socio- economic group of 95 school going children of Varanasi under an ICMR study, Department of Pediatrics, BHU during 1983-84 has been taken under study,(Khare and Sinha (2007)). The first 25% (i.e. 24 children) units have been considered as non-response units. The values of parameters related to the study characters y_1 (the height of children in cm.) and y_2 (weight of children in kg), the auxiliary character x (chest circumference of the children in cm) have been given as follows:

$\bar{Y}_1 = 115.9526$; $\bar{Y}_2 = 19.4968$

$\bar{X} = 55.8611$; $C_{y_1} = 0.05146$

$C_{y_2} = 0.15613$; $C_x = 0.05860$

$C_{y_{1(2)}}^{'} = 0.04402$; $C_{y_{2(2)}}^{'} = 0.12075$

$C_{x_{(2)}}^{'} = 0.05402$; $\rho_{01} = 0.620$

$\rho_{02} = 0.846$; $\rho_{12} = 0.713$

$\rho_{01}^{'} = 0.401$; $\rho_{02}^{'} = 0.729$

$\rho'_{12} = 0.678$

The problem considered is to estimate the ratio between height and weight of the male children aged 6-7 years using chest circumference as the auxiliary character.**Table 1**

Mean square error (MSE) and Relative efficiency (RE) of the estimators $\hat{R}^*$, $\hat{R}_1^*$ and $\hat{R}_a^*$ with respect to $\hat{R}^*$ for k=2, 3, 4 (N=95, n=35).

Estimator	$\frac{1}{k}$		
	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$
$\hat{R}^*$	100(0.010649185)	100(0.014690116)	100(0.0123130098)
$\hat{R}_1^*$	79(0.013478771)	119(0.012297874)	110(0.01116977)
$\hat{R}_a^*$	80(0.013274067)	120(0.012161457)	111(0.011048847)

Figures in parenthesis give the MSE.

Table 2

Relative efficiency (RE) of $\hat{R}^*$, $\hat{R}_1^*$ and $\hat{R}_a^*$ with respect to $\hat{R}^*$ for fixed $C_0 = 225$; $(c = Rs.2; c_1 = Rs.5; c_2 = Rs.15)$

Estimator	k_{opt}	n_{opt}	R.E
$\hat{R}^*$	1.9264620196	29.20596455	100(0.021492255)
$\hat{R}_1^*$	2.852967446	31.84668754	109(0.019593901)
$\hat{R}_a^*$	2.946104093	32.03709741	111(0.019385824)

Figures in parenthesis give the MSE.

Table 3

Expected cost of the estimators $\hat{R}^*$, $\hat{R}_1^*$ and $\hat{R}_a^*$ for the specified precision $V_0 = 0.674$;

$c = Rs.2; c_1 = Rs.5; c_2 = Rs.15$

Estimator	k_{opt}	n_{opt}	Expected cost (in Rs.)
$\hat{R}^*$	1.9264620196	0.923368172	7.174714504
$\hat{R}_1^*$	2.852967446	0.925817292	6.54099082
$\hat{R}_a^*$	2.946104093	0.921462253	6.471529062

From table 1, we observed that the estimators $\hat{R}_1^*$ and $\hat{R}_a^*$ are more efficient than the estimator $\hat{R}^*$ for $k = 2,3$. Also it is observed that the estimator $\hat{R}_a^*$ is more efficient than the estimator $\hat{R}^*$ and $\hat{R}_1^*$ for different values of k .

From table 2, we observed that for fixed cost the estimators $\hat{R}_1^*$ and $\hat{R}_a^*$ have smaller mean square error than that of the estimator $\hat{R}^*$ Also it is observed that the estimator $\hat{R}_a^*$ has the smallest mean square error for fixed cost corresponding to the estimators $\hat{R}^*$ and $\hat{R}_1^*$.

From table 3, we observed that expected cost incurred in the estimator $\hat{R}_a^*$ is less as compared to the expected cost incurred for $\hat{R}^*$ and $\hat{R}_1^*$ in the case of specified precision.

References

1. Hansen M.H., Hurwitz W.N. (1946): The problem of non-response in sample surveys; JASA, 41: 517- 529.
2. Khare BB, Srivastava S (1997): Transformed ratio type estimators for the population mean in the presence of non-response; Commun Stat Theory Methods 26(7):1779-1791.
3. Okafor FC, Lee H (2000): Double sampling for ratio and regression estimation with subsampling the non respondents; Surv Methodol 26(2):183-188.
4. Ray SK and Singh RK (1979): A generalized ratio - product - difference (RPD) estimator in double sampling; J Ind Soc Agril Stat XXXI(1): 85 - 88.
5. Rao PSRS (1986): Ratio estimation with sub-sampling the non-respondents; Surv Methodol 12(2):217-230. . Rao PSRS (1987): Ratio and regression estimates with sub-sampling the non-respondents. Paper presented at a special contributed session of the International Statistical Association Meeting, Sept., 2-16, Tokyo, Japan.
7. Singh R. Karan and Bhushan S. (2013): A Class of Estimators of Population Mean in Presence of Non Response, International Journal of Statistics and Analysis (Accepted).
8. Tabasum R, Khan IA (2004): Double sampling for ratio estimation with non-response; J Ind Soc Agril Stat 58(3):300-306

* * *

[1]Department of Applied Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow
shashi.py@gmail.com
[2]Department of Statistics, Lucknow University, Lucknow