International Journal of Applied Agricultural Research
ISSN 0973-2683 Volume 12, Number 1 (2017) pp. 21-32

http://www.ripublication.com

On a Class of Cost Efficient Estimators of Product of Population Means in Presence of Non Response

Shashi Bhushan, Praveen Kumar Misra and Nazia Naqvi

Department of Mathematics and Statistics,
Dr. Shakuntala Misra National Rehabilitation University, India

Abstract

In the presence of non-response, a class of estimators of product 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. An empirical study is also done.

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

1. INTRODUCTION

In surveys covering human populations, information is in most cases not obtainedfrom all the units in the survey even after some call-backs. An estimator based on such incomplete information is generally biased and the results may be grossly misleading when the respondents differ from the non-respondents. Hansen and Hurwitz(1946)considered the problem of non-response while estimating the population mean by taking a sub sample from the non-respondent group with the help of some extra efforts and an estimatorwas proposed by combining the information available from response and non-responsegroups. In this method, questionnaires are mailed to all the respondents included in a sample and a list of non-respondents is prepared after the deadline is over. Then a sub sample is drawn from the set of non-respondents and a direct interview is conducted with the selected respondents and the necessary information is collected.

In estimating population parameters like the mean, total or ratio, sometimes we make use of auxiliary information to improve precision of the estimates. The estimation of ratio of two population means usingauxiliary information was considered by Singh (1965), Tripathi (1970), Singh (1998), Upadhayay *et al.*(2000) and Birader and Singh (1997-1998). When the auxiliary information on a variable x is present in the form of its population mean $\bar{X}$, and the non-response is present, then the problem of estimation of ratio of two population means $\bar{Y}_1$ and $\bar{Y}_2$ of the studyvariates $\bar{y}_1$ and $\bar{y}_2$ has been dealt by various authors including Rao(1986, 1987) &Khare and Srivastava (1993).If the population mean $\bar{X}$ of the auxiliary variable x is not known in presence of non-response then we make use of double sampling (or two phase sampling) procedure to estimate 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 SRSWOR. Then a second phase sample of size $n\left(n < n^{'}\right)$ is drawn from $n^{'}$ by SRSWOR and the information on study variables is measured on it.

In this paper, instead of resorting to double sampling, if the population mean $\bar{X}$ of the auxiliary variable x is not known, here 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 SRSWOR assuming that the complete information on the auxiliary variable x is available to us. Let n_1 units provide information on study variables $\bar{y}_1$ and $\bar{y}_2$; and n_2 units do not respond. Then utilizing Hansen and Hurwitz technique we consider a sub sample from n_2 non-responding units of size $r(=\frac{n_2}{k}, k > 1)$. Now, our proposed class of estimators for estimating the product of the two population means is

$$\hat{P}_a^* = \hat{P}^* \left(\frac{\bar{x}^*}{\bar{x}} \right)^a$$

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 $\left(\bar{y}_{i(1)}, \bar{x}_{(1)}\right)$ and $\left(\bar{y}_{i(2)}, \bar{x}_{(2)}\right)$ 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 estimator $\hat{P}_a^*$

Let us define

$$e_1^* = \frac{\bar{y}_1^* - \bar{Y}_1}{\bar{Y}_1} \quad ; \quad e_2^* = \frac{\bar{y}_2^* - \bar{Y}_2}{\bar{Y}_2} \quad ; \quad e_0^* = \frac{\bar{x}^* - \bar{X}}{\bar{X}} \quad ; \quad e_0 = \frac{\bar{x} - \bar{X}}{\bar{X}}$$

so that $E\left(e_1^*\right)=E\left(e_2^*\right)=E\left(e_0^*\right)=E\left(e_0\right)=0$ and

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

$$E\left(e_0^{*2}\right)=\frac{(N-n)}{Nn}C_x^2+\frac{(k-1)N_2}{Nn}C_{x_{(2)}}^{*2};\ E\left(e_0^2\right)=\frac{(N-n)}{Nn}C_x^2$$

$$E\left(e_0^*e_1^*\right)=\frac{(N-n)}{Nn}\rho_{12}C_{y_1}C_{y_2}+\frac{(k-1)N_2}{Nn}\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*;\ E\left(e_1^*e_0\right)=\frac{(N-n)}{Nn}\rho_{01}C_{y_1}C_x$$

$$E\left(e_0^*e_1^*\right)=\frac{(N-n)}{Nn}\rho_{01}C_{y_1}C_x+\frac{(k-1)N_2}{Nn}\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^*;\ E\left(e_2^*e_0\right)=\frac{(N-n)}{Nn}\rho_{02}C_{y_2}C_x$$

$$E\left(e_0^*e_2^*\right)=\frac{(N-n)}{Nn}\rho_{02}C_{y_2}C_x+\frac{(k-1)N_2}{Nn}\rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*;\ E\left(e_0^*e_0\right)=\frac{(N-n)}{Nn}C_x^2$$

Now $\hat{P}_a^*=\hat{P}^*\left(\frac{\bar{x}^*}{\bar{x}}\right)^a=P(1+e_1^*)(1+e_2^*)(1+e_0^*)^a(1+e_0)^{-a}$

$$=P(1+e_1^*+e_2^*+e_1^*e_2^*)\left\{1-ae_0+\frac{a(a-1)}{2!}e_0^2+ae_0^*-a^2e_0^*e_0+\frac{a(a-1)}{2!}e_0^{*2}+\ldots\right\}$$

$$=P+P\left\{ae_0^*-ae_0+e_1^*+e_2^*+\frac{a(a-1)}{2!}e_0^{*2}+\frac{a(a-1)}{2!}e_0^2-a^2e_0^*e_0-ae_0e_2^*+ae_0^*e_2^*-ae_0e_1^*+ae_0^*e_1^*+e_1^*e_2^*+\ldots\right\}$$

Therefore

$$\hat{P}_a^*-P=P\left\{e_1^*+e_2^*+ae_0^*-ae_0+\frac{a(a-1)}{2!}e_0^{*2}+\frac{a(a-1)}{2!}e_0^2+e_1^*e_2^*-a^2e_0^*e_0-ae_0e_2^*+ae_0^*e_2^*-ae_0e_1^*+ae_0^*e_1^*+\ldots\right\}\quad(2.1)$$

Now, on taking expectation on both sides of (2.1) and ignoring the error terms of degree greater than two, we have bias of $\hat{P}_a^*$ to the first degree of approximation, that is, upto the terms of order $0\left(\frac{1}{n}\right)$ to be

$$Bias\left(\hat{P}_a^*\right)=P\left[\frac{(N-n)}{Nn}\left\{\rho_{12}C_{y_1}C_{y_2}-aC_x^2\right\}+\frac{(k-1)N_2}{Nn}\left\{\frac{a(a-1)}{2!}C_{x_{(2)}}^{*2}+\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right.\right.$$

$$\left.\left.+a\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^*+a\rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right\}\right]\quad(2.2)$$

Further, 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 $\hat{P}_a^*$ to the terms of order $0\left(\frac{1}{n}\right)$ to be

$$MSE(\hat{P}_a^*)=P^2\left[\left\{\frac{(N-n)}{Nn}\left(C_{y_1}^2+C_{y_2}^2+2\rho_{12}C_{y_1}C_{y_2}\right)+\frac{(k-1)N_2}{Nn}\left(C_{y_{1(2)}}^{*2}+C_{y_{2(2)}}^{*2}+2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right)\right\}\right.$$

$$\left.+\frac{(k-1)N_2}{Nn}\left\{a^2C_{x_{(2)}}^{*2}+2a\left(\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^*+\rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right)\right\}\right]\quad(2.3)$$

$$MSE(\hat{P}_a^*) = MSE(\hat{P}) + \frac{(k-1)N_2}{Nn} P^2 \left[a^2 C_{x_{(2)}}^{*2} + 2a\left(\rho_{01}^* C_{y_{1(2)}}^* C_{x_{(2)}}^* + \rho_{02}^* C_{y_{2(2)}}^* C_{x_{(2)}}^* \right) \right] \tag{2.4}$$

From (2.4), the optimum value of *a* minimizing the mean square error of the estimator is

$$a_{opt.} = -\left\{ \rho_{01}^* \left(\frac{C_{y_{1(2)}}^*}{C_{x_{(2)}}^*} \right) + \rho_{02}^* \left(\frac{C_{y_{2(2)}}^*}{C_{x_{(2)}}^*} \right) \right\} \tag{2.5}$$

and the minimum mean square error is given by

$$Min\, MSE(\hat{P}_a^*) = P^2 \left[\left\{ \frac{(N-n)}{Nn} \left(C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12} C_{y_1} C_{y_2} \right) + \frac{(k-1)N_2}{Nn} \left(C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} + 2\rho_{12}^* C_{y_{1(2)}}^* C_{y_{2(2)}}^* \right) \right\} - \frac{(k-1)N_2}{Nn} \left(\rho_{01}^* C_{y_{1(2)}}^* + \rho_{02}^* C_{y_{2(2)}}^* \right)^2 \right]$$

$$Min\, MSE(\hat{P}_a^*) = MSE(\hat{P}^*) - P^2 \frac{(k-1)N_2}{Nn} \left(\rho_{01}^* C_{y_{1(2)}}^* + \rho_{02}^* C_{y_{2(2)}}^* \right)^2 \tag{2.6}$$

showing that, for optimum value of*a*, the estimator is more efficient than the usual ratio estimator $\hat{P}^*$. The value $a_{opt.}$ 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

$$a_{opt} = -\left\{ \rho_{01}^* \left(\frac{C_{y_{1(2)}}^*}{C_{x_{(2)}}^*} \right) + \rho_{02}^* \left(\frac{C_{y_{2(2)}}^*}{C_{x_{(2)}}^*} \right) \right\} = -\frac{\left\{ \rho_{01}^* \frac{S_{y_{1(2)}}^* S_{x_{(2)}}^*}{\bar{Y}_1 \bar{X}} - \rho_{02}^* \frac{S_{y_{2(2)}}^* S_{x_{(2)}}^*}{\bar{Y}_2 \bar{X}} \right\}}{\frac{S_{x_{(2)}}^{*2}}{\bar{X}^2}} = -\frac{\left\{ \frac{S_{y_{1(2)}x_{(2)}}^*}{\bar{Y}_1 \bar{X}} - \frac{S_{y_{2(2)}x_{(2)}}^*}{\bar{Y}_2 \bar{X}} \right\}}{\frac{S_{x_{(2)}}^{*2}}{\bar{X}^2}} \tag{2.7}$$

The value $a_{opt.}$ 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 valueof *a* by estimated optimum c (based on sample observations) of a_{opt} given by

$$\hat{a}_{opt} = -\frac{\left\{ \frac{s_{y_{1(2)}x_{(2)}}^*}{\bar{x}^* \bar{y}_1^*} - \frac{s_{y_{2(2)}x_{(2)}}^*}{\bar{x}^* \bar{y}_2^*} \right\}}{\frac{s_{x_{(2)}}^{*2}}{\bar{x}^{*2}}} = -\frac{\left\{ \frac{r_{01}^* s_{y_{1(2)}}^* s_{x_{(2)}}^*}{\bar{x}^* \bar{y}_1^*} - \frac{r_{02}^* s_{y_{2(2)}}^* s_{x_{(2)}}^*}{\bar{x}^* \bar{y}_2^*} \right\}}{\frac{s_{x_{(2)}}^{*2}}{\bar{x}^{*2}}} \tag{2.8}$$

$$\hat{a}_{opt}=-\frac{\left\{\dfrac{r_{01}^{*}s_{y_{1(2)}}^{*}s_{x_{(2)}}^{*}}{\bar{x}^{*}\bar{y}_{1}^{*}}-\dfrac{r_{02}^{*}s_{y_{2(2)}}^{*}s_{x_{(2)}}^{*}}{\bar{x}^{*}\bar{y}_{2}^{*}}\right\}}{\dfrac{s_{x_{(2)}}^{*2}}{\bar{x}^{*2}}}=-\left\{\frac{\bar{x}^{*}s_{y_{1(2)}x_{(2)}}^{*}}{\bar{y}_{1}^{*}s_{x_{(2)}}^{*2}}+\frac{\bar{x}^{*}s_{y_{2(2)}x_{(2)}}^{*}}{\bar{y}_{2}^{*}s_{x_{(2)}}^{*2}}\right\}=C \qquad (2.9)$$

Replacing a_{opt} of a by the estimated optimum $\hat{C}$ in $\hat{P}_{a}^{*}$, we get the estimator $\hat{P}_{c}^{*}$ whose bias and mean square error are found below in the next section 3.

3. Bias and Mean Square Error of the estimator $\hat{P}_{c}^{*}$

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

$$e_3^{*}=\frac{\left(s'_{y_{1(2)}x_{(2)}}-S'_{y_{1(2)}x_{(2)}}\right)}{S'_{y_{1(2)}x_{(2)}}};\; e_4^{*}=\frac{\left(s'_{y_{2(2)}x_{(2)}}-S'_{y_{2(2)}x_{(2)}}\right)}{S'_{y_{2(2)}x_{(2)}}};\; e_5^{*}=\frac{s'^{2}_{x_{(2)}}-S'^{2}_{x_{(2)}}}{S'^{2}_{x_{(2)}}}$$

Now, we have

$$\hat{P}_{c}^{*}=\hat{P}^{*}\left(\frac{\bar{x}^{*}}{\bar{x}}\right)^{\hat{a}_{opt}}=P(1+e_1^{*})(1+e_2^{*})\left\{\frac{(1+e_0^{*})\bar{X}}{(1+e_0)\bar{X}}\right\}^{\hat{a}_{opt}}$$

$$=P+P\left[\left(e_1^{*}+e_2^{*}\right)+\hat{a}_{opt}\left(e_0^{*}-e_0-e_0e_2^{*}+e_0^{*}e_2^{*}-e_0e_1^{*}+e_0^{*}e_1^{*}\right)+\frac{\hat{a}_{opt}\left(\hat{a}_{opt}-1\right)}{2!}\left(e_0^{*2}+e_0^{2}\right)-\hat{a}_{opt}^{2}e_0^{*}e_0+e_1^{*}e_2^{*}+\ldots\right]$$

Therefore

$$\hat{P}_{c}^{*}-P=P\left[\left(e_1^{*}+e_2^{*}+e_1^{*}e_2^{*}\right)+\left\{\frac{\bar{X}S_{y_{1(2)}x_{(2)}}^{*}}{\bar{Y}_{1}S_{x_{(2)}}^{*2}}+\frac{\bar{X}S_{y_{2(2)}x_{(2)}}^{*}}{\bar{Y}_{2}S_{x_{(2)}}^{*2}}\right\}\left(e_0-e_0^{*}+e_0e_1^{*}-e_0^{*}e_1^{*}+\ldots\right)\right] \qquad (3.1)$$

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 $\hat{P}_{c}^{*}$ to the first degree of approximation is of order $0\left(\frac{1}{n}\right)$; 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 $\hat{P}_{c}^{*}$ to the first degree of approximation, that is, up to terms of order $0\left(\frac{1}{n}\right)$ is

$$MSE(\hat{P}_c^*) = P^2\left[\left\{\frac{(N-n)}{Nn}\left(C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12}C_{y_1}C_{y_2}\right) + \frac{(k-1)N_2}{Nn}\left(C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} + 2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right)\right\} - \frac{(k-1)N_2}{Nn}\left(\rho_{01}^*C_{y_{1(2)}}^* + \rho_{02}^*C_{y_{2(2)}}^*\right)^2\right]$$

$$\mathrm{MSE}(\hat{\mathrm{P}}_c^*) = \mathrm{MSE}(\hat{\mathrm{P}}^*) - \mathrm{P}^2\,\frac{(\mathrm{k}-1)\mathrm{N}_2}{\mathrm{Nn}}\left(\rho_{01}^*\mathrm{C}_{y_{1(2)}}^* + \rho_{02}^*\mathrm{C}_{y_{2(2)}}^*\right)^2 \qquad (3.2)$$

Thus, we see that the estimated optimum mean square error is equal to the minimum mean square error.

4. Determination of n and k for fixed precision

Let us consider a cost function for $\hat{P}_a^*$ as, $C = cn + c_1n_1 + c_2r$

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_1n and $\frac{W_2n}{k}$. Thus the expected cost is given by, $E(C) = C^* = n\left[c + c_1W_1 + \frac{c_2W_2}{k}\right]$

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

$$= n\left[c + c_1W_1 + \frac{c_2W_2}{k}\right] + \lambda\left\{P^2\left[\left\{\frac{(N-n)}{Nn}\left(C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12}C_{y_1}C_{y_2}\right)\right.\right.\right.$$

$$\left. + \frac{(k-1)N_2}{Nn}\left(C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} - 2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right)\right\} + \frac{(k-1)N_2}{Nn}\left\{a^2C_{x_{(2)}}^{*2}\right.$$

$$\left.\left.\left. + 2a\left(\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^* + \rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right)\right\}\right] - V_0\right\}$$

$$= \mathrm{n}\left[\mathrm{c} + \mathrm{c}_1\mathrm{W}_1 + \frac{\mathrm{c}_2\mathrm{W}_2}{\mathrm{k}}\right] + \lambda\left\{\left[\frac{(\mathrm{N}-\mathrm{n})}{\mathrm{Nn}}\mathrm{U}_1 + \frac{(\mathrm{k}-1)\mathrm{N}_2}{\mathrm{Nn}}\mathrm{U}_2\right]\right\} \qquad (4.2)$$

where $U_1 = \left\{C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12}C_{y_1}C_{y_2}\right\}$

$$U_2 = \left\{C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} + 2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right\} + a^2C_{x_{(2)}}^{*2} + 2a\left\{\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^* + \rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right\}$$

and λ is Lagrange's multiplier.

Now differentiating (4.2) with respect to n and k , and on equating them with zero we get $n = \sqrt{\frac{\lambda P^2\left\{U_1 + (k-1)W_2U_2\right\}}{\left\{c + c_1W_1 + \frac{c_2W_2}{k}\right\}}}$ $\quad \frac{n}{k} = \sqrt{\frac{\lambda P^2U_2}{c_2}}$ (4.4)

On putting (4.3) in (4.4) we get, $k_{opt} = \sqrt{\dfrac{c_2(U_1 - W_2U_2)}{(c + c_1W_1)U_2}}$ (4.5)

which is required optimum value of k. Further substituting the value of n and k in the expression on of MSE we get.

$$\sqrt{\lambda} = \frac{\sqrt{\left\{U_1 + (k-1)W_2U_2\right\}\left\{c + c_1W_1 + \dfrac{c_2W_2}{k}\right\}}}{\left\{\dfrac{V_0}{P} + \dfrac{PU_1}{N}\right\}} \qquad (4.6)$$

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

$$n_{opt} = \frac{\sqrt{\left\{U_1 + (k_{opt.} - 1)W_2U_2\right\}}}{\left\{\dfrac{V_0}{P^2} + \dfrac{U_1}{N}\right\}} \qquad (4.7)$$

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

$$C^* = \frac{\left\{U_1 + (k_{opt} - 1)W_2U_2\right\}\left\{c + c_1W_1 + \dfrac{c_2W_2}{k_{opt}}\right\}}{\left\{\dfrac{V_0}{P^2} + \dfrac{U_1}{N}\right\}} \qquad (4.8)$$

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

$$C^* = \frac{P^2\left\{U_1 + (k_{opt} - 1)W_2U_2\right\}\left\{c + c_1W_1 + \dfrac{c_2W_2}{k_{opt}}\right\}}{V_0}$$

$$C^* = \frac{P^2}{V_0}\Big[\left\{C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12}C_{y_1}C_{y_2}\right\} + (k_{opt} - 1)W_2\Big\{\left(C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} + 2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right) + a^2C_{x_{(2)}}^{*2}$$

$$+2a\left(\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^* + \rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right)\Big\}\Big]\left[c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right] \qquad (4.9)$$

5. Determination of Optimum Values of n & 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 $C = n\left(c + c_1W_1 + \dfrac{c_2W_2}{k}\right)$

(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 $\hat{P}_a^*$ can be expressed as, $MSE\left(\hat{P}_a^*\right) = P^2\left[\frac{U_1}{n} + \frac{kW_2U_2}{n} - \frac{W_2U_2}{n} - \frac{U_1}{N}\right]$ (5.2)

To determine the optimum values of n and k that minimize the $MSE(\hat{P}_a^*)$ for a fixed cost($C^* < C_0$) , we consider the function, $\phi^* = MSE\left(\hat{P}_a^*\right) + \lambda\left\{\left(c + c_1W_1 + \frac{c_2W_2}{k}\right)\right\}$ where λ is Lagrange's multiplier. (5.3)

Now differentiating (5.3) with respect to n and k ; and on equating them with zero we get,

$$n = \sqrt{\frac{P^2\left\{U_1 + (k-1)W_2U_2\right\}}{\lambda\left\{c + c_1W_1 + \frac{c_2W_2}{k}\right\}}} \tag{5.4}$$

$$\frac{n}{k} = \sqrt{\frac{P^2U_2}{\lambda c_2}} \tag{5.5}$$

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

$$k_{opt} = \sqrt{\frac{c_2\left(U_1 - W_2U_2\right)}{\left(c + c_1W_1\right)U_2}} \tag{5.6}$$

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

$$\sqrt{\lambda} = \sqrt{\frac{P^2\left\{U_1 + \left(k_{opt.} - 1\right)W_2U_2\right\}\left\{c + c_1W_1 + \frac{c_2W_2}{k_{opt.}}\right\}}{C}} \tag{5.7}$$

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

$$n_{opt} = \frac{C}{\left\{c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right\}} \tag{5.8}$$

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

$$MSE\left(\hat{P}_a^*\right) = P^2\left[\frac{\left\{U_1 + \left(k_{opt}-1\right)W_2U_2\right\}\left\{c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right\}}{C} - \frac{U_1}{N}\right] \tag{5.9}$$

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

$$MSE\left(\hat{P}_a^*\right) = \frac{P^2\left\{U_1 + \left(k_{opt}-1\right)W_2U_2\right\}\left\{c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right\}}{C}$$

Therefore, for the optimum values of n and k, the minimum mean square error of $\hat{P}_a^*$ for fixed cost ($C \le C_0$) is given as,

$$MSE\left(\hat{P}_a^*\right) = \frac{P^2}{C}\left[\left\{C_{y_1}^2 + C_{y_2}^2 + 2\rho_{12}C_{y_1}C_{y_2}\right\} + (k_{opt}-1)W_2\left\{\left(C_{y_{1(2)}}^{*2} + C_{y_{2(2)}}^{*2} + 2\rho_{12}^*C_{y_{1(2)}}^*C_{y_{2(2)}}^*\right)\right.\right.$$
$$\left.\left. + a^2C_{x_{(2)}}^{*2} + 2a\left(\rho_{01}^*C_{y_{1(2)}}^*C_{x_{(2)}}^* + \rho_{02}^*C_{y_{2(2)}}^*C_{x_{(2)}}^*\right)\right\}\right]\left[c + c_1W_1 + \frac{c_2W_2}{k_{opt}}\right] \tag{5.10}$$

6. 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:

$\bar{Y}_1 = 115.9526;\ \bar{Y}_2 = 19.4968;\ \bar{X} = 55.8611;$

$C_{y_1} = 0.05146;\ C_{y_2} = 0.15613;\ C_x = 0.05860;$

$C_{y_{1(2)}}^* = 0.04402;\ C_{y_{2(2)}}^* = 0.12075;\ C_{x_{(2)}}^* = 0.05402;$

$C_{x_{(2)}}^* = 0.05402;\ \rho_{02} = 0.846;\ \rho_{12} = 0.713;$

$\rho_{01}^* = 0.401;\ \rho_{02}^* = 0.729;\ \rho_{12}^* = 0.678;$

The problem considered is to estimate the product 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 Percent relative efficiency (PRE) of the estimators $\hat{P}^*$, $\hat{P}_1^*$ and $\hat{P}_a^*$ with respect to $\hat{P}^*$ for *k*=2, 3, 4 (*N*=95, *n*=35).

Estimator	$1/k$		
	$1/2$	$1/3$	$1/4$
$\hat{P}^*$	100(4424.226)	100(5299.476)	100(6174.726)
$\hat{P}_1^*$	108(4110.685)	113 (4672.394)	118(5234.104)
$\hat{P}_a^*$	110(4012.24)	118(4475.504)	125(4938.768)

Figures in parenthesis give the MSE.

Table 2

Percent relative efficiency (RE) of $\hat{P}^*$, $\hat{P}_1^*$ and $\hat{P}_a^*$ with respect to $\hat{P}^*$ for fixed $C_0 = 315$; $(c = Rs.2.8; c_1 = Rs.7; c_2 = Rs.21)$

Estimator	k_{opt}	n_{opt}	P.R.E.
$\hat{P}^*$	1.892158	29.07139	100(7705.258)
$\hat{P}_1^*$	2.438738	30.8612	106 (7289.339)
$\hat{P}_a^*$	2.711396	31.53711	108 (7116.105)

Figures in parenthesis give the MSE.

Table 3

Expected cost of the estimators $\hat{P}^*$, $\hat{P}_1^*$ and $\hat{P}_a^*$ for the specified precision $V_0 = 0.9436$; $c = Rs.2.8; c_1 = Rs.7; c_2 = Rs.21$

Estimator	k_{opt}	n_{opt}	Expected Cost(inRs.)
$\hat{P}^*$	1.892158	108.1522	2572230
$\hat{P}_1^*$	2.438738	108.6134	2433385
$\hat{P}_a^*$	2.711396	108.3544	2375554

From table 1, we observed that the estimators $\hat{P}_1^*$ and $\hat{P}_a^*$ are more efficient than the estimator $\hat{P}^*$ for $k = 2, 3, 4$. Also it is observed that the estimator $\hat{P}_a^*$ is more efficient than the estimator $\hat{P}^*$ and $\hat{P}_1^*$ for different values of k.

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

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

7. Conclusion

From the preceding sections it is ample clear that the proposed classes are an improvement over the conventional estimators under non-response.

REFERENCES

[1] Bhushan S. and Naqvi N. (2015) "Generalized efficient classes of estimators in presence of non-response using two auxiliary variables", Journal of Statistics and Management System, (ISSN: 0972-0510), 18(6), 573-602.

[2] Shashi Bhushan and Nazia Naqvi (2015) "A cost efficient estimator in presence of non-response for ratio of two population means using two auxiliary variables", Proceeding of National Conference on Recent Advances in Statistical and Mathematical Sciences and their applications (RAMSA 2014), Kumaun University, 125 – 133.

[3] Shashi Bhushan , R. Karan Singh and Nazia Naqvi (2013) "A class of cost efficient estimators of population ratio of means in presence of non response", Mathematical Sciences International Research Journal (ISSN: 2278 – 8697), 2(2), 662 – 668.

[4] Shashi Bhushan (2013)"A cost efficient class of estimators of population mean in presence of non- response when the population mean of auxiliary variable is not known without using double sampling", International Journal of Multidisciplinary Research (ISSN: 2277 – 9302), 12(1), 32 – 35.

[5] Hansen M.H., Hurwitz W.N. (1946): The problem of non-response in sample surveys; JASA,41: 517- 529.

[6] Khare BB, Srivastava S (1997): Transformed ratio type estimators for the population mean inthe presence of non-response; Commun Stat Theory Methods 26(7):1779-1791.

[7] Okafor FC, Lee H (2000): Double sampling for ratio and regression estimation with subsampling the non-respondents; SurvMethodol 26(2):183-188.

[8] Ray SK and Singh RK (1979): A generalized ratio - product - difference (RPD) estimator in double sampling; J IndSocAgril Stat XXXI(1): 85 - 88.

[9] Rao PSRS (1986): Ratio estimation with sub-sampling the non-respondents; SurvMethodol 12(2):217-230.

[10] 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.

[11] 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).

[12] Tabasum R, Khan IA (2004): Double sampling for ratio estimation with non-response; J IndSocAgril Stat 58(3):300-306.