J. Stat. Appl. Pro. Lett. **2**, No. 3, 263-273 (2015)

Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.12785/jsapl/020308

A Class of Estimators for Population Mean Utilizing the Information on Auxiliary Variable and the Attribute using Two Phase Sampling Scheme in the Presence of Non-Response

B. B. Khare[*] and *P. S. Jha*

Department of Statistics, Banaras Hindu University, Varanasi -221005, India.

Received: 17 Nov. 2014, Revised: 27 Mar.2015, Accepted: 28Mar. 2015.
Published online: 1Sep. 2015.

Abstract: In this paper, we have proposed two classes of estimators for population mean in which first is using an attribute and second is using an auxiliary variable and an attribute both with unknown population means using two-phase sampling scheme in the presence of non- response on the study variable. The properties of the proposed class of estimators have been studied for a fixed sample size, for a fixed cost and also for a specified precision. The empirical study has shown that the proposed class of estimators using auxiliary variable and attribute are of great importance in increasing the efficiency and decreasing the cost of the survey.

Keywords: Non Response, Auxiliary Variable, Attribute, population Mean, Two Phase.

1 Introduction

Sometimes, it may not be possible to collect the complete information for all the units selected in the sample due to lack of interest, person not present at home, lack of knowledge regarding the survey, ethical problems and refusal of the respondent to respond for the given questionnaire. For estimation of population mean in sample surveys when a sample of size n is drawn from the population of size N by using SRSWOR method of sampling, then it has been observed that n_1 units respond and n_2 units do not respond. Further, a sub-sample of size $r(r = n_2 k^{-1}, k > 1)$ from n_2 non-responding units has been drawn by using SRSWOR method of sampling by making extra effort and proposed the estimator of population mean by Hansen and Hurwitz [12] which is given as follow:

$$\bar{y}^* = \frac{n_1}{n}\bar{y}_1 + \frac{n_2}{n}\bar{y}_2' \ , \tag{1}$$

where $\bar{y}_1$ and $\bar{y}_2'$ are sample mean based on n_1 responding units and sub sampled $r\,(\frac{n_2}{k}, k > 1)$ units from non-responding units. Similarly in case of auxiliary variable and attribute, we have

$$\bar{x}^* = \frac{n_1}{n}\bar{x}_1 + \frac{n_2}{n}\bar{x}_2' \ , \tag{2}$$

$$\bar{\phi}^* = \frac{n_1}{n}\bar{\phi}_1 + \frac{n_2}{n}\bar{\phi}_2' \ , \tag{3}$$

where $\bar{x}_1$ and $\bar{x}_2'$ are sample mean based on corresponding n_1 responding units and r sub sampled units from n_2 non responding units on the auxiliary variable x and $\bar{\phi}_1$ and $\bar{\phi}_2'$ are sample proportion based on corresponding n_1 responding units and r sub sampled units from n_2 non responding units on the attribute ϕ. Sometimes auxiliary informations are available in the form of attribute. In this case Naik and Gupta [18], Jhajj et.al [9], Shabbir and Gupta [10,11] Singh et.al [17], Koyuncu [13] and Singh and Solanki [8] have proposed different type of estimators for population mean when population proportion $\bar{\phi}_N$ for attribute ϕ is known and unknown. These estimators are proposed in the case of complete

[*]Corresponding author e-mail: bbkhare56@yahoo.com

response on the units selected in the sample for the study variable and the attribute.

In the case when $\bar{X}$ is known, the conventional and alternate ratio, product, regression and generalized estimators have been proposed by Rao [14, 15], Khare and Srivastava [3, 5].In the case, the two-phase sampling ratio, product and regression type estimators for population mean $\bar{Y}$, using auxiliary variable in the presence of non-response have been proposed by Khare [1] and Khare and Srivastava [3, 4, 6].

Further the class of estimator for population mean in the presence of non-response on the study variable y using two-phase sampling scheme with auxiliary variable have been proposed by Khare and Sinha [2] and Khare et.al [7].

In this paper, we have proposed two classes of estimators using an attribute with unknown value of the population proportion using two-phase sampling scheme in the presence of non- response. Further, we have extended it to another class of estimators for population mean in the form of a general class of estimators for $\bar{Y}$ using an auxiliary variable and an attribute with both unknown values of the population mean of the auxiliary variable and the population proportion using two-phase sampling scheme in the presence of non- response on the study variable. An empirical study has been conducted and a comparative study of the proposed class of estimators has been made with the relevant estimators. The properties of the proposed class of estimators have been studied for fixed sample sizes (n', n), for a fixed cost and also for a specified precision.

2Proposed Estimators and Their Biases and Mean Square Errors (MSE)

Let us consider a finite population $U = \{U_i, i = 1,2,...,N\}$ of N identifiable units. Let y_i, x_i and ϕ_i denote the value of i_{th} units for the study variable y, auxiliary variable x and the attribute ϕ. Here it is to be noted that ϕ_i takes value 1 if it belong to attribute ϕ otherwise zero. Here we are considering the problem of estimation of the population mean $\bar{Y}$ of study variable in the presence of non-response on the study variable y when $\bar{X}$ the population mean of the auxiliary variable and the population proportion $\bar{\phi}_N$ of the attribute ϕ are not known. In the first phase, we select a sample of size n' from the population of size N by using SRSWOR sampling scheme to estimate $\bar{X}$ and $\bar{\phi}_N$. Again we select a second phase sample of size n from n' by using SRSWOR sampling scheme to obtain the required information on the study variable y and $\bar{y}^*$, $\bar{x}^*$ and $\bar{\phi}^*$ are defined in eq (1),(2) and (3) of section 1.

Following Khare and Sinha [2], Sinha and Kumar [16], we have suggested improved classes of estimators for estimating population mean of the study variable using auxiliary information in the form of auxiliary variable and attribute with known and unknown population means in the presence of non-response on the study variable only.

Further, we propose a class of estimators for $\bar{Y}$ using attribute ϕ in the presence of non response which is given by

$$T_1 = f(w, v), \tag{4}$$

Such that $f(\bar{Y},1) = \bar{Y}$ and $f_1(\bar{Y},1) = \left(\dfrac{\partial f}{\partial \bar{y}^*}\right)_{(\bar{Y},1)} = 1$,

Where $w = \bar{y}^*$, $v = \dfrac{\bar{\phi}^*}{\bar{\phi}'_n}$ and $\bar{\phi}'_n = \dfrac{1}{n'}\sum_{1}^{n'} \phi_i$.

The function $f(w, v)$ satisfies the following regularity conditions:

(i) For any sampling design whatever be the sample chosen, the function $f(w, v)$ assumes values in a bounded, closed convex subset D of the two dimensional real space containing the point $(\bar{Y},1)$. (5)

(ii) The function $f(w,v)$ and its first and second order partial derivative exist and are continuous and bounded in D. (6)

Expanding the function $f(w,v)$ about the point $(\bar{Y},1)$ by using Taylor's series upto the second order partial derivatives and using the regularity conditions, we have

$$T_1=\bar{Y}+(w-\bar{Y})f_1(a)+(v-1)f_2(a)+\frac{1}{2}\{(w-\bar{Y})^2 f_{11}(a^*)+(v-1)^2 f_{22}(a^*)+2(w-\bar{Y})(v-1)f_{12}(a^*)\} \qquad (7)$$

Where $f(a)=\bar{Y}$, $f_1(a)=1$, $f_{11}(a)=0$, $a=(\bar{Y},1)$, $a^*=(w^*,v^*)$, $w^*=\bar{Y}+\theta_1(w-\bar{Y})$ and $v^*=1+\theta_2(v-1)$, $0<\theta_i<1 \;\; \forall\; i=1,2$.

For the large sample approximation, the expressions for bias and mean square error (MSE) of T_1 upto the terms of order (n^{-1}) are given by

$$\text{Bias}(T_1)=\left\{\frac{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')}{\bar{\phi}_N}\right\}f_{12}(a^*) \qquad (8)$$

and

$$\begin{aligned}\text{MSE}(T_1)&=E\{(w-\bar{Y})^2+(v-1)^2 f_2^{\,2}(a)+2(w-\bar{Y})(v-1)f_2(a)\}\\&=V(\bar{y}^*)+\frac{V(\bar{\phi}^*)-V(\bar{\phi}')}{\bar{\phi}_N^2}f_2^{\,2}(a)+2\frac{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')}{\bar{\phi}_N}f_2(a)\end{aligned} \qquad (9)$$

Differentiating $\text{MSE}(T_1)$ w.r.to $f_2(a)$ and equating to zero, we get

$$f_2(a)_{opt}=-\bar{\phi}_N\left\{\frac{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')}{V(\bar{\phi}^*)-V(\bar{\phi}')}\right\} \qquad (10)$$

And the minimum value of $MSE(T_1)$ is given by

$$MSE(T_1)_{\min.}=V(\bar{y}^*)-\frac{\{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')\}^2}{V(\bar{\phi}^*)-V(\bar{\phi}')}, \qquad (11)$$

where $V(\bar{y}^*)=\frac{N-n}{Nn}S_y^2+\frac{W_2(k-1)}{n}S_{y_2}^2$, $V(\bar{\phi}^*)=\frac{N-n}{Nn}S_\phi^2+\frac{W_2(k-1)}{n}S_{\phi_2}^2$, $V(\bar{\phi}')=\frac{N-n'}{Nn'}S_\phi^2$,

$S_\phi^2=\frac{1}{N-1}\sum_{i=1}^{N}(\phi_i-\bar{\phi}_N)^2$, $S_{\phi_2}^2=\frac{1}{N_2-1}\sum_{i=1}^{N_2}(\phi_i-\bar{\phi}_{N_2})^2$, $Cov(\bar{y}^*,\bar{\phi}^*)=\frac{N-n}{Nn}S_{y\phi}+\frac{W_2(k-1)}{n}S_{y\phi_{(2)}}$,

$Cov(\bar{y}^*,\bar{\phi}')=\frac{N-n'}{Nn'}S_{y\phi}$, $S_{y\phi}=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})(\phi_i-\bar{\phi}_N)$, $S_{y\phi_{(2)}}=\frac{1}{N_2-1}\sum_{i=1}^{N_2}(y_i-\bar{Y}_2)(\phi_i-\bar{\phi}_{N_2})$.

Putting $S_{y\phi}=S_{y\phi_{(2)}}$ and $S_{\phi}^2=S_{\phi_2}^2$ in the equation (11) the expression for the $MSE(T_1)_{\min.}$ is given by

$$MSE(T_1)_{\min.}=\frac{1}{n}\{(S_y^2+W_2(k-1)S_{y_2}^2)-((1+W_2(k-1))\frac{(S_{y\phi})^2}{S_{\phi}^2})\}+\frac{1}{n'}\frac{(S_{y\phi})^2}{S_{\phi}^2}+\frac{1}{N}S_y^2. \tag{12}$$

Some members of the class T_1 are given as follow:

$$T_{11}=\bar{y}^*(a_1+(1-a_1)v),\ T_{12}=\bar{y}^*(v^{\alpha}),$$

$$T_{13}=(\bar{y}^*+a_2(v-1))v^{\beta},\ T_{14}=\bar{y}^*(2-v^{\alpha'}),$$

$$T_{15}=\bar{y}^*\exp(\frac{v-1}{v+1})\text{ and }T_{16}=\bar{y}^*(b_0+(1-b_0)v^{\alpha}).$$

Again using auxiliary variable x and the attribute ϕ, we propose a class of estimators T_2 in the presence of non-response on the study variable which is given as follow:

$$T_2=h(w,u,v), \tag{13}$$

Such that $h(\bar{Y},1,1)=\bar{Y}$ and $h_1(\bar{Y},1,1)=\left(\frac{\partial h}{\partial \bar{y}^*}\right)_{(\bar{Y},1,1)}=1$,

Where $w=\bar{y}^*$, $u=\frac{\bar{x}^*}{\bar{x}'}$ and $v=\frac{\bar{\phi}^*}{\bar{\phi}_n'}$.

Let function $h(w,u,v)$ satisfies the following regularity conditions:

(iii) For any sampling design whatever be the sample chosen the function $h(w,u,v)$ assume values in a bounded, closed convex subset D_1 of the three dimensional real space containing the point $(\bar{Y},1,1)$. (14)

(iv) The function $h(w,u,v)$ and its first and second order partial derivative exist and are continuous and bounded in D_1. (15)

Now, expanding the function $h(w,u,v)$ about the point $(\bar{Y},1,1)$ by using Taylor's series up to the second order partial derivatives and using the regularity conditions, we have

$$\begin{aligned}T_2=&\bar{Y}+(w-\bar{Y})h_1(G)+(u-1)h_2(G)+(v-1)h_3(G)+\frac{1}{2}\{(w-\bar{Y})^2h_{11}(G^*)+(u-1)^2h_{22}(G^*)+(v-1)^2h_{33}(G^*)\}\\&+\{(w-\bar{Y})(u-1)h_{12}(G^*)+(u-1)(v-1)h_{23}(G^*)+(w-\bar{Y})(v-1)h_{13}(G^*)\}\end{aligned} \tag{16}$$

where $h(G)=\bar{Y}$, $h_1(G)=1$, $h_{11}(G)=0$, $G=(\bar{Y},1,1)$, $G^*=(w^*,u^*,v^*)$, $w^*=\bar{Y}+\theta_1(w-\bar{Y})$ $u^*=1+\theta_2(u-1)$ and $v^*=1+\theta_3(v-1)$, $0<\theta_i<1\ \ \forall\ i=1,2,3$.

For the large sample approximation, the expressions for bias and mean square error (MSE) of T_2 upto the terms of order (n^{-1}) are given by

$$\text{Bias}(T_2)=\frac{Cov(\bar{y}^*,\bar{x}^*)-Cov(\bar{y}^*,\bar{x}')}{\bar{X}}h_{12}(G^*)+\frac{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')}{\bar{\phi}_N}h_{13}(G^*)$$
$$+\frac{\left\{Cov(\bar{x}^*,\bar{\phi}^*)-Cov(\bar{x}^*,\bar{\phi}')\right\}}{\bar{X}\bar{\phi}_N}h_{23}(G^*) \quad (17)$$

and

$$\text{MSE}(T_2)=E\{(w-\bar{Y})^2+(u-1)^2h_2^2(G)+(v-1)^2h_3^2(G)+2(w-\bar{Y})(u-1)h_2(G)$$
$$+2(w-\bar{Y})(v-1)h_3(G)+2(u-1)(v-1)h_2(G)h_3(G)\} \quad (18)$$

Differentiating $\text{MSE}(T_1)$ w.r.to $h_2(G)$, $h_3(G)$ and equating to zero, we get the optimum value of $h_2(G)$ and $h_3(G)$ which are given as follows:

$$h_2(G)_{opt}=\frac{D_1E_1-B_1C_1}{A_1C_1-E_1^2}\text{ and }h_3(G)_{opt}=-\frac{(E_1h_2(G)+D_1)}{C_1}. \quad (19)$$

Now, putting the optimum value of $h_2(G)_{opt}$ and $h_3(G)_{opt}$ in the $MSE(T_2)$, we get

$$MSE(T_2)_{\text{min.}}=V(\bar{y}^*)-\frac{\{D_1\}^2}{C_1}-\frac{(D_1E_1-B_1C_1)^2}{C_1(A_1C_1-E_1^2)}, \quad (20)$$

where $A_1=\dfrac{V(\bar{x}^*)-V(\bar{x}')}{\bar{X}^2}$, $B_1=\dfrac{Cov(\bar{y}^*,\bar{x}^*)-Cov(\bar{y}^*,\bar{x}')}{\bar{X}}$,

$D_1=\dfrac{Cov(\bar{y}^*,\bar{\phi}^*)-Cov(\bar{y}^*,\bar{\phi}')}{\bar{\phi}_N}$, $C_1=\dfrac{V(\bar{\phi}^*)-V(\bar{\phi}')}{\bar{\phi}_N^2}$ and

$$E_1=\frac{Cov(\bar{x}^*,\bar{\phi}^*)-Cov(\bar{x}^*,\bar{\phi}')+Cov(\bar{x}',\bar{\phi}')-Cov(\bar{x}',\bar{\phi}^*)}{\bar{X}\bar{\phi}_N}=\frac{Cov(\bar{x}^*,\bar{\phi}^*)-Cov(\bar{x}',\bar{\phi}^*)}{\bar{X}\bar{\phi}_N}.$$

For $S_{y\phi}=S_{y\phi_{(2)}}$, $S_{yx}=S_{yx_{(2)}}$, $S_{x\phi}=S_{x\phi_{(2)}}$, $S_x^2=S_{x_2}^2$ and $S_\phi^2=S_{\phi_2}^2$ in the equation (20) we get the $MSE(T_2)_{\text{min.}}$ which is given as follow:

$$MSE(T_2)_{\text{min.}}=\frac{1}{n}[S_y^2+W_2(k-1)S_{y_2}^2-(1+W_2(k-1))\{\frac{(S_{y\phi})^2}{S_\phi^2}+\frac{((S_{y\phi}S_{x\phi})-(S_{yx}S_\phi^2))^2}{S_\phi^2((S_x^2S_\phi^2)-(S_{x\phi})^2)}\}]$$
$$+\frac{1}{n'}\{\frac{(S_{y\phi})^2}{S_\phi^2}+\frac{((S_{y\phi}S_{x\phi})-(S_{yx}S_\phi^2))^2}{S_\phi^2((S_x^2S_\phi^2)-(S_{x\phi})^2)}\}+\frac{1}{N}S_y^2. \quad (21)$$

Some member of the class T_2:

$T_{21}=\bar{y}^*u^{\alpha_1}v^{\alpha_2}$, $T_{22}=\bar{y}^*(bu^{\alpha_3}+(1-b)v^{\alpha_4})$ and $T_{23}=\bar{y}^*(1+a_5(u-1)+\alpha_6(v-1))$,

where $\alpha_i \quad \forall\ i=1,2,3,4,5,6$ and $\alpha,\beta,a_1,a_2,b,b_0$ are constant.

3 Comparison of the Proposed Estimators with the Relevant Estimators

On comparing $\bar{y}^*$, T_1 and T_2, we observe that .

$$V(\bar{y}^*) - MSE(T_1)_{\min.} = \frac{\{Cov(\bar{y}^*, \bar{\phi}^*) - Cov(\bar{y}^*, \bar{\phi}')\}^2}{V(\bar{\phi}^*) - V(\bar{\phi}')} \geq 0, \tag{22}$$

$$V(\bar{y}^*) - MSE(T_2)_{\min.} = \frac{\{D_1\}^2}{C_1} + \frac{(D_1E_1 - B_1C_1)^2}{C_1(A_1C_1 - E_1^2)} \geq 0, \tag{23}$$

$$MSE(T_1)_{\min.} - MSE(T_2)_{\min.} = \frac{(D_1E_1 - B_1C_1)^2}{C_1(A_1C_1 - E_1^2)} \geq 0. \tag{24}$$

Hence we see that

$$MSE(T_2)_{\min.} \leq MSE(T_1)_{\min.} \leq V(\bar{y}^*). \tag{25}$$

4 Determination of n', n and k for a Fixed Cost ($C \leq C_0$)

Let C_0 be the total cost (fixed) of the survey apart from the overhead cost. The cost function C' for the cost incurred in the survey apart from overhead expenses, can be expressed by

$$C' = C_1'n' + C_1 n + C_2 n_1 + C_3 \frac{n_2}{k}. \tag{26}$$

Since C' will vary from sample to sample, so we consider the expected cost C to be incurred in the survey apart from overhead expenses, which is given by

$$C = E(C') = C_1'n' + n\left[C_1 + C_2W_1 + C_3\frac{W_2}{k}\right], \tag{27}$$

Where

C_1' : The cost per unit of identifying and observing attribute / auxiliary variable or both,

C_1 : The cost per unit of mailing questionnaire/visiting the unit at the second phase,

C_2 : The cost per unit of collecting and processing data for the study variable y obtained from n_1 responding units and

C_3 : The cost per unit of obtaining and processing data for the study variable y (after extra efforts) from sub-sampled units.

From equation (12) and (21) the $MSE(T_i)_{\min}, i = 1,2$; can be expressed in terms of the notation Ψ_{ni}, $\Psi_{n'i}$ and $\Psi_{n/ki}$ which is given as

$$MSE(T_i) = \frac{1}{n'}\Psi_{n'i} + \frac{1}{n}\Psi_{ni} + \frac{k}{n}\Psi_{n/ki} + \text{terms of independent of } n' \text{ and } n, \tag{28}$$

Where $\Psi_{n'i} =$ coefficient of $\frac{1}{n'}$ terms, $\Psi_{ni} =$ coefficient of $\frac{1}{n}$ terms and $\Psi_{n/ki} =$ coefficient of $\frac{k}{n}$ terms.

Now we minimize MSE for the fixed cost $C \leq C_0$ and to obtain the optimum values of n', n and k.

Let us consider a function

$$\Omega = MSE(T_i) + \lambda_i \left\{ C_1' n' + n\left(C_1 + C_2 W_1 + C_3 \frac{W_2}{k} \right) - C_0 \right\},\ i = 1,2 \tag{29}$$

where λ_i are Lagrange's multipliers.

Now differentiating Ω with respect to n', n and k and equating to zero, we get

$$n'_{opt} = \sqrt{\frac{\Psi_{n'i}}{\lambda_i C_1'}}, \tag{30}$$

$$n_{opt} = \sqrt{\frac{\Psi_{ni} + k_{opt}\Psi_{n/ki}}{\lambda_i \left(C_1 + C_2 W_1 + C_3 \frac{W_2}{k_{opt}} \right)}}, \tag{31}$$

$$\text{and } \frac{n}{k} = \sqrt{\frac{\Psi_{n/ki}}{\lambda_i C_3 W_2}}. \tag{32}$$

Solving (31) and (32), we have

$$k_{opt.} = \sqrt{\frac{\Psi_{ni} C_3 W_2}{\Psi_{n/ki}(C_1 + C_2 W_1)}}. \tag{33}$$

Putting the values of n', n and k in (29), we have

$$\sqrt{\lambda_i} = \frac{1}{C_0}\left[\sqrt{\Psi_{n'i} C_1'} + \sqrt{(\Psi_{ni} + k_{opt}\Psi_{n/ki})(C_1 + C_2 W_1 + C_3 \frac{W_2}{k_{opt.}})} \right]. \tag{34}$$

The minimum value of $MSE(T_i),\ i = 1,2$ is given by

$$MSE(T_i)_{\min.} = \frac{1}{C_0}\left[\sqrt{\Psi_{n'i} C_1'} + \sqrt{(\Psi_{ni} + k_{opt.}\Psi_{n/ki})(C_1 + C_2 W_1 + C_3 \frac{W_2}{k_{opt.}})} \right]^2 - \frac{S_y^2}{N}. \tag{35}$$

In case of $\bar{y}^*$, the expected total cost is given by

$$C = E(C') = n\left(C_1 + C_2 W_1 + C_3 \frac{W_2}{k} \right) \tag{36}$$

$$\text{And} \qquad MSE(\bar{y}^*)_{\min.} = \frac{1}{C_0}\left[\sqrt{(S_y^2 + W_2(k-1)S_{y_2}^2)(C_1 + C_2 W_1 + C_3 \frac{W_2}{k})} \right]^2 - \frac{S_y^2}{N}. \tag{37}$$

5 Determination of n', n and k for a Specified Variance

Let V_0' be the variance of the estimator $T_i, (i=1,2)$ fixed in advance and we have

$$V_0' = \frac{1}{n}\Psi_{ni} + \frac{1}{n'}\Psi_{n'i} + \frac{k}{n}\Psi_{n/ki} - \frac{S_y^2}{N}. \quad (38)$$

For minimizing the average total cost C for the specified variance of the estimator T_i, we define a function

$$\Omega_1 = \left\{ C_1'n' + n\left(C_1 + C_2W_1 + C_3\frac{W_2}{k} \right) \right\} + \mu_i\,\{MSE(T_i) - V_0'\}. \quad (39)$$

Where $\mu_i\ (i=1,2)$ are Lagrange multiplier.

Now differentiating with respect to n', n and k and equating to zero, we have

$$n'_{opt} = \sqrt{\frac{\mu_i \Psi_{n'i}}{C_1'}}, \quad (40)$$

$$n_{opt} = \sqrt{\frac{\mu_i(\Psi_{ni} + k_{opt}\Psi_{n/ki})}{\left(C_1 + C_2W_1 + C_3\frac{W_2}{k_{opt.}} \right)}}, \quad (41)$$

$$\text{And } \frac{n}{k} = \sqrt{\frac{\mu_i \Omega_{n/ki}}{C_3W_2}}. \quad (42)$$

From (39), (41) and (42), we have

$$k_{opt.} = \sqrt{\frac{\Psi_{ni}C_3W_2}{\Psi_{n/ki}(C_1 + C_2W_1)}}, \quad (43)$$

$$\text{and } \sqrt{\mu_i} = \frac{1}{\left[V_0' + \frac{S_y^2}{N} \right]}\left[\sqrt{C_1'\Psi_{n'i}} + \sqrt{(\Psi_{ni} + k_{opt.}\Psi_{n/ki})\left(C_1 + C_2W_1 + C_3\frac{W_2}{k_{opt.}} \right)} \right]. \quad (44)$$

The minimum expected total cost for the specified variance V_0' for the estimator T_i is given by

$$C(T_i)_{\min.} = \frac{\left[\sqrt{C_1'\Psi_{n'i}} + \sqrt{(\Psi_{ni} + k_{opt.}\Psi_{n/ki})\left(C_1 + C_2W_1 + C_3\frac{W_2}{k_{opt.}} \right)} \right]^2}{\left[V_0' + \frac{S_y^2}{N} \right]},\ i=1,2. \quad (45)$$

In case of $\bar{y}^*$, the fixed precision is given by

$$V_0 = \frac{1}{n}\Psi_{ni} + \frac{k}{n}\Psi_{n/ki} - \frac{S_y^2}{N} \quad (46)$$

$$C(\bar{y}^*)_{\min.} = \frac{\left[\sqrt{(\Psi_{ni} + k_{opt.}\Psi_{n/ki})\left(C_1 + C_2 W_1 + C_3 \frac{W_2}{k_{opt.}}\right)}\right]^2}{\left[V_0 + \frac{S_y^2}{N}\right]}, \tag{47}$$

6 An Empirical Study

To study the performance of the proposed classes of estimators, we have considered the population of 128 villages from District census Handbook (1951 Census) for Madurai District, Madras State, and the total number of persons, total area (in sq. miles) and cultivated area (in acres) for village was considered. The first 25% villages (32) have been considered as non-response group of the population. Here we have taken the study variable y, auxiliary variable x and attribute ϕ which are given as follow:

y = Total cultivated area (in acres) in a village.

x = Total number of persons in a village.

$$\phi = \begin{cases} 1, \text{ if total area (in sq. miles) greater than 5.6 for a village} \\ 0, \text{ otherwise.} \end{cases}$$

The values of the parameters of the population under study are given as follows:

$$\bar{Y} = 1943.375,\ \bar{X} = 3243.352,\ P = 0.414,\ C_y = 0.572,\ C_x = 0.605,\ C_p = 1.194,$$

$$\bar{Y}_2 = 1984.5,\ \bar{X}_2 = 2790.594,\ P_2 = 0.182,\ C_{y(2)} = 0.632,\ C_{x(2)} = 0.555,\ C_{p(2)} = 0.84,$$

$$\rho_{yx} = 0.492,\ \rho_{yx(2)} = 0.812,\ \rho_{py} = 0.653,\ \rho_{py(2)} = 0.786,\ \rho_{px} = 0.469,\ \rho_{px(2)} = 0.766.$$

Here the problem is to estimate the total cultivated area (in acres) in a village by using total number of persons in a village as the auxiliary variable and total area (in sq. miles) greater than 5.6 in a village as attribute. The value of $f_2(a)$, $h_2(G)$ and $h_3(G)$ from equation (10) and (19) can be calculated and these values are given as follow:

$$f_2(a)_{opt_{(k=2)}} = -702.040,\quad f_2(a)_{opt_{(k=3)}} = -787.923,\ f_2(a)_{opt_{(k=4)}} = -866.622,$$

$$h_2(G)_{opt_{(k=2)}} = -781.515,\quad h_2(G)_{opt_{(k=3)}} = -985.636,\ h_2(G)_{opt_{(k=4)}} = -1119.075,$$

$$h_3(G)_{opt_{(k=2)}} = -489.462,\quad h_3(G)_{opt_{(k=3)}} = -488.909 \text{ and } h_3(G)_{opt_{(k=4)}} = -494.964.$$

Table1: Relative efficiency (RE in %) of the proposed estimator with respect to relevant estimator.

Estimators	$N = 128,\ n' = 64,\ n = 32$ RE (in %)		
	$1/k$		
	1/2	1/3	1/4
$\bar{y}^*$	100.00 (41250.48)#	100.00 (53539.76)	100.00 (65829.04)
$T_{1\min.}$	138.67 (29746.69)	139.47 (38387.80)	140.96 (46698.99)
$T_{2\min.}$	157.32 (26220.33)	169.68 (31552.58)	181.03 (36362.14)

(# Figures in parenthesis show the mean square error of the estimators)

From the table 1, we observed that the classes estimator T_1 and T_2 are more efficient than $\bar{y}^*$ for different values of k. MSE of T_1 and T_2 decreases on k decreases, but it is observed that the MSE of T_2 decreases with the faster rate in comparison to the MSE of T_1. Further it is observed that RE of T_1 and T_2 increases as sub sampling fraction decreases. It is also observed that there is highly increases in RE of T_2 than RE of T_1 as k decreases, which shows that the addition of an auxiliary information in T_2 has greater role in the estimator in increasing the efficiency in comparison to T_1.

Table2: Relative efficiency and expected cost of the proposed classes of estimators for fixed cost as well as for fixed precision.

$C_1' = Rs.5$, $C_1 = Rs.15$, $C_2 = Rs.75$, $C_3 = Rs.1500$								
Estimators	$C_0 =$ Rs. 3000 (fixed)					$V_0' = 50000$ acre^2 (fixed)		
	$k_{opt.}$ (~)	$n'_{opt.}$ (~)	$n_{opt.}$ (~)	$MSE(.)$ (~)	R.E. in % w.r.t. $\bar{y}^*$	$n'_{opt.}$ (~)	$n_{opt.}$ (~)	Expected cost in Rs.
$\bar{y}^*$	3.36	_	21	95196.79	100.00	_	36	5272.96
T_1	3.00	62	17	71540.64	133.07	85	23	4083.28
T_2	2.94	67	16	67732.15	140.55	87	21	3891.75

From the table 2, we observe that for fixed cost, $MSE(T_2)_{\min.} \le MSE(T_1)_{\min.} \le V(\bar{y}^*)$ and also for fixed precision the cost of the estimator T_2 is minimum.

7 Conclusion

So, we observe that $MSE(T_2)_{\min.} \le MSE(T_1)_{\min.} \le V(\bar{y}^*)$ for fixed sample size (n', n) and for fixed cost. It is also observed that the cost incurred for T_2 is less than the cost incurred for T_1 and $\bar{y}^*$. Hence the use of an auxiliary variable and an attribute together along with the study variable in the estimation of population mean in increasing the efficiency and decreasing the cost of the survey is more useful than the estimators using only an attribute along with the study variable.

References

[1] B. B. Khare, Non-response and the estimation of population mean in its presence of sample survey. Seminar proceeding. 12th orientation course (Science Stream) sponsored by U.G.C. Academic Staff College, B.H.U., India (1992).

[2] B. B. Khare, R. R. Sinha, General class of two phase sampling estimators for the population mean using an auxiliary character in the presence of non-response. Proceedings of the 5th International Symposium on Optimization and Statistics, 233-245 (2002).

[3] B. B. Khare, S. Srivastava, Estimation of population mean using auxiliary character in presence of non-response. Nat. Acad. Sc. Letters, India **16**, 111-114 (1993).

[4] B. B. Khare, S. Srivastava, Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of non-response, Proc. Nat. Acad. Sci, India, **65(A) II**, 195-203 (1995).

[5] B. B. Khare, S. Srivastava, Generalised estimators for population mean in presence of nonresponse. Internal. J. Math. & Statist. Sci. **9**, 75-87 (2000).

[6] B. B. Khare, S. Srivastava, Generalized two phase estimators for the population mean in the presence of non-response. Ali. Jour. Stat. **30**, 39-54 (2010).

[7] B. B. Khare, U. Srivastava, K. Kumar, An Improved class of two phase sampling estimator for population mean using auxiliary character in the presence of non response. J. Sc. Res., BHU, Vol. **55**, 151-161 (2011).

[8] H. P. Singh, R. S. Solanki, Improved estimation of population mean in simple random sampling using information on auxiliary attribute. Appl. Math. Comp. **218**, 7798-7812 (2012).

[9] H. S. Jhajj, M. K. Sharma, L. K. Grover, A family of estimators of population mean using auxiliary attribute, Pak. J. Stat. **22 (1)**, 43-50 (2006).

[10] J. Shabbir, S. Gupta, On estimating the finite population mean with known population proportion of an auxiliary variable, Pak. J. Stat. **23 (1)**, 1-9 (2007).

[11] J. Shabbir, S. Gupta, Estimation of the finite population mean in two phase sampling when auxiliary variables are attributes. Hacett. J. Math. Stat. **39(1)**, 121-129 (2010).

[12] M. H. Hansen, W. N. Hurwitz, The problem of non-response in sample surveys. J. Amer. Stat. Assoc. **41**, 517-529 (1946).

[13] N. Koyuncu, Efficient estimators of population mean using auxiliary attributes. Appl. Math. Comp. **218,** 10900-10905 (2012).

[14] P. S. R. S. Rao, Ratio estimation with subsampling the nonrespondents. Survey Methodology **12**, 217-230 (1986).

[15] P. S. R. S. Rao, Regression estimators with subsampling of nonrespondents. In-Data Quality Control, Theory and Pragmatics, (Eds.) Gunar E. Liepins and V.R.R. Uppuluri, Marcel Dekker, New York, 191-208 (1990).

[16] R. R. Sinha, V. Kumar, improved classes of estimators for population mean using information on auxiliary character under double sampling the non-respondents, Nat. Acad. Sc. Letters, India **37(1)**, 71-79 (2014).

[17] R. Singh, P. Chauhan, N. Sawan, F. Smarandache, Ratio estimators in simple random sampling using information on auxiliary attribute. Pak. J. Stat. and Operation Research. **4**, 47-53 (2008).

[18] V. D. Naik, P. C. Gupta, A note on estimation of means with known population proportion of an auxiliary character, J. Ind. Soc. Agri. Stat. **48(2)**, 151-158 (1996).

B. B. Khare, presently working as Professor of Statistics, Department of Statistics, Facultyof Science, Banaras Hindu University, Varanasi, India since 2006. Joined the Department as lecturer in 1984. Dr. Khare did his M.Sc. (Statistics) and Ph.D. (Statistics) from Banaras Hindu University in 1977 and 1984 respectively. He has more than 30 years teaching and research experience. The fields of specialization are Sampling Theory and Population studies (physicalgrowth and development, breast feeding and weaning practices during infancy), published 105 research papers in the reputed international/national journals. Guided 5 Ph. D. scholars for their Ph. D. Degree. About 60 papers presented in international/national conferences, Life member of National Academy of Sciences and M. N. A. Sc. in India and reviewer of a number of international/national journals of repute

Prem Shankar Jha, pursuing research work for Ph.D. degree in Statistics in the Department of Statistics, Banaras Hindu University Varanasi-221005, India, on the topic " Estimation of Population Parameter's Using Auxiliary Variable in Stratified Population in the Presence of Non-Response". He has published 6 research papers in the reputed international/national journals. He has done B.Sc. and M. Sc. in Statistics from Department of Statistics, Banaras Hindu University Varanasi-221005, India,