Utilization of Some Known Population Parameters for Estimating Population Mean in Presence of Non-Response

Sunil Kumar
Department of Statistics
University of Jammu
Jammu and Kashmir, India
sunilbhougal06@gmail.com

Abstract

In presence of non-response, a general class of estimators is proposed for estimating the population mean of the study variable with known value of some population parameter(s) of the auxiliary variable. The bias and mean squared error (MSE) of the proposed strategy are derived under simple random sampling without replacement (SRSWOR). Some estimators are also obtained from proposed strategy by giving suitable values to the constants used. Comparisons of the proposed strategy with the usual unbiased estimator and other estimators have been made. An empirical study is carried out to demonstrate the performance of the proposed estimator.

Keywords: Bias, Mean Squared Error (MSE), Non-response.

2000 AMS Classification: 62D05.

1. Introduction

In many human surveys, information is in most cases not obtained from all the units in the survey even after some call-backs. Hansen and Hurwitz (1946) have given a sampling plan which calls for taking a sub sample of non-respondents after the first mail attempt and then enumerating the sub sample by personal interviews. Later, the problem of non-response has already been tackled from various angles by various authors among El-Badry (1956), Cochran (1977), Rao (1986, 1987), Khare and Srivastava (1993, 1995, 1997), Okafor and Lee (2000), Tabasum and Khan (2004, 2006), Sodipo and Obisesan (2007) and Singh and Kumar (2008 a, b, 2009, 2010 a, b, 2011).

Let y and x be the study and the auxiliary variable under study with population mean $\bar{Y}$ and $\bar{X}$ respectively on a finite population $U = (U_1, U_2, ..., U_N)$. Consider a simple random sample of size n drawn without replacement from N units. When non-response occurs in the initial attempt, Hansen and Hurwitz (1946) suggested a double sampling scheme for estimating the populationmean comprising the following steps:

(a) a simple random sample of size n is selected without replacement and the questionnaire is mailed to the sampled units:

(b) a sub sample of size $r = n_2/k$, $(k \geq 1)$ from the n_2 non-responding units in the initial attempt is contacted through personal interviews.

An unbiased estimator $\bar{y}^*$ for the population mean $\bar{Y}$ of the study variable y is defined by

$$\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_{2r},$$

where $w_1 = n_1/n$, $w_2 = n_2/n$, $\bar{y}_1 = \sum_{i=1}^{n_1} y_i/n_1$ and $\bar{y}_{2r} = \sum_{i=1}^{r} y_i/r$.

The variance of $\bar{y}^*$ is given by

$$Var(\bar{y}^*) = \left(\frac{1-f}{n}\right)S_y^2 + \frac{W_2(k-1)}{n}S_{y(2)}^2,$$

where $W_2 = N_2/N$, $f = n/N$, $S_y^2 = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1)$,

$S_{y(2)}^2 = \sum_{i=1}^{N_2}(y_i - \bar{Y}_2)^2/(N_2-1)$, $\bar{Y} = \sum_{i=1}^{N} y_i/N$, $\bar{Y}_2 = \sum_{i=1}^{N_2} y_i/N_2$, N_1 and $N_2(= N - N_1)$ are the sizes of the responding and non-responding units from the finite population N.

It is well known that in estimating population mean, sample survey experts sometimes use auxiliary information to improve precision of the estimates. Let x denote an auxiliary variate with population mean $\bar{X} = \sum_{i=1}^{N} x_i/N$. Let $\bar{X}_1 = \sum_{i=1}^{N_1} x_i/N_1$ and $\bar{X}_2 = \sum_{i=1}^{N_2} x_i/N_2$ denote the means of the response and non-response groups (or strata). Let $\bar{x} = \sum_{i=1}^{n} x_i/n$ denote the mean of all the n units. Let $\bar{x}_1 = \sum_{i=1}^{n_1} x_i/n_1$ and $\bar{x}_2 = \sum_{i=1}^{n_2} x_i/n_2$ denote the means of the n_1 responding units and n_2 non-responding units. Further, let $\bar{x}_{2r} = \sum_{i=1}^{r} x_i/r$ denote the mean of the $r(= n_2/k), k > 1$ sub-sampled units. With this background we define an unbiased estimator of population mean $\bar{X}$ as

$$\bar{x}^* = w_1\bar{x}_1 + w_2\bar{x}_{2r}.$$

The variance of $\bar{x}^*$ is given by

$$Var(\bar{x}^*) = \left(\frac{1-f}{n}\right)S_x^2 + \frac{W_2(k-1)}{n}S_{x(2)}^2,$$

where

$S_x^2 = \sum_{i=1}^{N}(x_i - \bar{X})^2/(N-1)$, $S_{x(2)}^2 = \sum_{i=1}^{N_2}(x_i - \bar{X}_2)^2/(N_2-1)$ and $\bar{X} = \sum_{i=1}^{N} y_i/N$.

In this paper, we have suggested a general family of estimators for estimating the population mean $\bar{Y}$ of the study variable y in presence of non-response by using some prior values of certain population parameters (i.e. Coefficient of variation (C_x), Correlation coefficient (ρ_{yx}), etc.). Expressions of bias and mean squared error (MSE), up to first degree of approximation, have been obtained, which will enable us to obtain the said expressions for any member of this family. Some well used estimators have been shown as the particular member of this family. The comparison is carried out between the suggested estimator with well-known estimators. An empirical study is also carried out to demonstrate the performance of the proposed class of estimators.

2. The proposed family of estimators

We consider the situation in which information on the auxiliary variable x is obtained from all the sample units (i.e. the initial sample units) and the population mean $\bar{X}$ of the auxiliary variable is known, but some sample units fail to supply information on the study variable y. While suggesting the estimator for the population mean $\bar{Y}$, Rao (1986) used only the information on the sample mean $\bar{x}$ and on the population mean $\bar{X}$ of the auxiliary variable x. However, one can also obtain the unbiased estimator $\bar{x}^* = w_1\bar{x}_1 + w_2\bar{x}_{2r}$ of $\bar{X}$ (without any extra effort) while in the process of obtaining $\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_{2r}$, the unbiased estimator of the population mean $\bar{Y}$. Thus in the present situation, we have two unbiased estimators, $\bar{x}^*$ and $\bar{x}$, of the population mean $\bar{X}$ of the auxiliary variable x. With this background, we define the following estimator of the population mean $\bar{Y}$ of the study variable y as

$$T = \bar{y}^*\left(\frac{a\bar{x}^* + b}{a\bar{X} + b}\right)^{\alpha}\left(\frac{a\bar{x} + b}{a\bar{X} + b}\right)^{\beta}, \qquad (2.1)$$

where $a(\neq 0)$ and b are either real numbers or the functions of the known population parameters such as standard deviation (σ), Coefficient of variation (C_x), Correlation coefficient (ρ_{yx}) etc. of the auxiliary variable x and (α, β) are suitable chosen constants.

To obtain the bias and variance of the class of estimators T, we write

$\bar{y}^* = \bar{Y}(1+\varepsilon_0)$, $\bar{x}^* = \bar{X}(1+\varepsilon_1)$, $\bar{x} = \bar{X}(1+\varepsilon_2)$

such that

$E(\varepsilon_0) = E(\varepsilon_1) = E(\varepsilon_2) = 0$, and

$E(\varepsilon_0^2) = \lambda C_y^2 + \lambda^* C_{y(2)}^2$, $E(\varepsilon_1^2) = \lambda C_x^2 + \lambda^* C_{x(2)}^2$, $E(\varepsilon_2^2) = \lambda C_x^2$,

$E(\varepsilon_0\varepsilon_1) = \lambda\rho_{yx}C_yC_x + \lambda^*\rho_{yx(2)}C_{y(2)}C_{x(2)}$, $E(\varepsilon_0\varepsilon_2) = \lambda\rho_{yx}C_yC_x$, $E(\varepsilon_1\varepsilon_2) = \lambda C_x^2$,

where

$\lambda = \left(\frac{1-f}{n}\right)$ and $\lambda^* = \frac{W_2(k-1)}{n}$.

Now expressing T in terms of $\varepsilon's$ we have

$$T = \overline{Y}(1+\varepsilon_0)(1+\phi\varepsilon_1)^{\alpha}(1+\phi\varepsilon_2)^{\beta}, \tag{2.2}$$

where $\phi = \left(\dfrac{a\overline{X}}{a\overline{X}+b}\right)$.

We assume that $|\phi\varepsilon_1| < 1$ and $|\phi\varepsilon_2| < 1$ so that the right hand side of (2.2) is expandable. Now, expanding the right hand side of (2.2) to the first degree of approximation, we have

$$\left(T-\overline{Y}\right) = \overline{Y}\left\{\varepsilon_0 + \alpha\phi\left(1+\varepsilon_0+\frac{\alpha-1}{2}\phi\varepsilon_1\right)\varepsilon_1 + \beta\phi\left(1+\varepsilon_0+\frac{\beta-1}{2}\phi\varepsilon_2\right)\varepsilon_2 + \alpha\beta\phi^{2}\varepsilon_1\varepsilon_2\right\} \tag{2.3}$$

Taking expectations of both sides of (2.3), we get the bias of T to the first degree of approximation is given by

$$B(T) = \overline{Y}\left[\lambda\phi\left\{\alpha\left(K_{yx}+\frac{\alpha-1}{2}\phi\right)+\beta\left(K_{yx}+\alpha\phi+\frac{\beta-1}{2}\phi\right)\right\}C_x^2 + \lambda^{*}\alpha\phi\left(K_{yx(2)}+\frac{\alpha-1}{2}\phi\right)C_{x(2)}^2\right], \tag{2.4}$$

where

$R = \overline{Y}/\overline{X}$, $K_{yx} = \beta_{yx}/R, K_{yx(2)} = \beta_{yx(2)}/R$, $\beta_{yx} = S_{yx}/S_x^2, \beta_{yx(2)} = S_{yx(2)}/S_{x(2)}^2$,

$C_x^2 = S_x^2/\overline{X}^2$, $C_y^2 = S_y^2/\overline{Y}^2, C_{x(2)}^2 = S_{x(2)}^2/\overline{X}^2$, $C_{y(2)}^2 = S_{y(2)}^2/\overline{Y}^2$,

$S_{yx} = \sum_{i=1}^{N}\left(x_i-\overline{X}\right)\left(y_i-\overline{Y}\right)/(N-1)$, $S_{yx(2)} = \sum_{i=1}^{N_2}\left(x_i-\overline{X}_2\right)\left(y_i-\overline{Y}_2\right)/(N_2-1)$, ρ_{yx}^2 and $\rho_{yx(2)}^2$ are respectively the correlation coefficient of response and non-response group between study variable y and auxiliary variable x and (α and β) are suitably chosen constants.

Squaring both sides of (2.3) and neglecting terms of $\varepsilon's$ involving power greater than two, we have

$$\left(T-\overline{Y}\right)^2 = \overline{Y}^2\left\{\varepsilon_0^2 + \alpha^2\phi^2\varepsilon_1^2 + \beta^2\phi^2\varepsilon_2^2 + 2\alpha\phi\varepsilon_0\varepsilon_1 + 2\alpha\beta\phi^2\varepsilon_1\varepsilon_2 + 2\beta\phi\varepsilon_0\varepsilon_2\right\}. \tag{2.5}$$

Taking expectations of both sides of (2.5), we get the variance of T to the first degree of approximation, we get

$$Var(T) = \overline{Y}^2\left[\lambda\left\{C_y^2 + (\alpha+\beta)\phi\left((\alpha+\beta)\phi+2K_{yx}\right)C_x^2\right\} + \lambda^{*}\left\{C_{y(2)}^2 + \alpha\phi\left(\alpha\phi+2K_{yx(2)}\right)C_{x(2)}^2\right\}\right], \tag{2.6}$$

The variance of T is minimized for

$$\alpha = -K_{yx(2)}/\phi \text{ and } \beta = (1/\phi)\left(K_{yx(2)}-K_{yx}\right) = -(1/\phi)\left(K_{yx}-K_{yx(2)}\right), \tag{2.7}$$

Putting (2.7) in (2.1), we get the asymptotically optimum estimator (AOE) as

$$T_{opt} = \overline{y}^{*}\left\{\frac{\left(a\overline{X}+b\right)^2}{\left(a\overline{x}^{*}+b\right)\left(a\overline{x}+b\right)}\right\}^{K_{yx(2)}/\phi}\left(\frac{a\overline{X}+b}{a\overline{x}+b}\right)^{K_{yx}/\phi}. \tag{2.8}$$

The minimum variance of T is given by

$$\min .Var(T)=\bar{Y}^{2}\left[\lambda\left(1-\rho_{yx}^{2}\right)C_{y}^{2}+\lambda^{*}\left(1-\rho_{yx(2)}^{2}\right)C_{y(2)}^{2}\right]=Var\left(T_{opt}\right), \qquad (2.9)$$

3. Some members of the proposed class of estimators T

The following are the estimators of the population mean which can be obtained bysuitable choices of constants α,β,a and b.

Estimator	α	β	a	b
$T_0=\bar{y}^*$ Usual unbiased estimator	0	0	a	b
$T_1=\bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right)$ Rao (1986) estimator	-1	0	1	0
$T_2=\bar{y}^*\left(\frac{\bar{X}}{\bar{x}}\right)$ Rao (1986) estimator	0	-1	1	0
$T_3=\bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right)\left(\frac{\bar{X}}{\bar{x}}\right)$ Singh and Kumar (2008b) estimator	-1	-1	1	0
$T_4=\bar{y}^*\left(\frac{\bar{x}^*}{\bar{X}}\right)\left(\frac{\bar{x}}{\bar{X}}\right)$ Singh and Kumar (2008b) estimator	1	1	1	0
$T_5=\bar{y}^*\left(\frac{\bar{X}+C_x}{\bar{x}^*+C_x}\right)$	-1	0	1	C_x
$T_6=\bar{y}^*\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right)$	0	-1	1	C_x
$T_7=\bar{y}^*\left(\frac{\bar{X}+C_x}{\bar{x}^*+C_x}\right)\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right)$	-1	-1	1	C_x
$T_8=\bar{y}^*\left(\frac{\bar{X}+\rho_{yx}}{\bar{x}^*+\rho_{yx}}\right)$	-1	0	1	ρ_{yx}
$T_9=\bar{y}^*\left(\frac{\bar{X}+\rho_{yx}}{\bar{x}+\rho_{yx}}\right)$	0	-1	1	ρ_{yx}
$T_{10}=\bar{y}^*\left(\frac{\bar{X}+\rho_{yx}}{\bar{x}^*+\rho_{yx}}\right)\left(\frac{\bar{X}+\rho_{yx}}{\bar{x}+\rho_{yx}}\right)$	-1	-1	1	ρ_{yx}

where C_x is the coefficient of variation of the auxiliary variable x and ρ_{yx} is the correlation coefficient between the study variable y and the auxiliary variable x.

Many more estimators can also be generated from the proposed estimator at (2.1) just by putting different values of α, β, a and b.

The expressions of bias and Variance of the above estimators can be obtained by mere substituting the values of α, β, a and b in (2.4) and (2.6) respectively.

Up to the first degree of approximation, the Bias and Variance expressions of various estimators are

$$B(T_0) = \bar{Y}\ , \tag{3.1}$$

$$B(T_1) = \bar{Y}\left\{\lambda(1-K_{yx})C_x^2 + \lambda^*(1-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.2}$$

$$B(T_2) = \bar{Y}\lambda(1-K_{yx})C_x^2\ , \tag{3.3}$$

$$B(T_3) = \bar{Y}\left\{\lambda(3-2K_{yx})C_x^2 + \lambda^*(1-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.4}$$

$$B(T_4) = \bar{Y}\left\{\lambda(1+2K_{yx})C_x^2 + \lambda^* K_{yx(2)}C_{x(2)}^2\right\}, \tag{3.5}$$

$$B(T_5) = \bar{Y}\left\{\lambda\phi'(\phi'-K_{yx})C_x^2 + \lambda^*\phi'(\phi'-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.6}$$

$$B(T_6) = \bar{Y}\lambda\phi'(\phi'-K_{yx})C_x^2\ , \tag{3.7}$$

$$B(T_7) = \bar{Y}\left\{\lambda\phi'(3\phi'-2K_{yx})C_x^2 + \lambda^*\phi'(\phi'-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.8}$$

$$B(T_8) = \bar{Y}\left\{\lambda\phi^*(\phi^*-K_{yx})C_x^2 + \lambda^*\phi^*(\phi^*-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.9}$$

$$B(T_9) = \bar{Y}\lambda\phi^*(\phi^*-K_{yx})C_x^2\ , \tag{3.10}$$

$$B(T_{10}) = \bar{Y}\left\{\lambda\phi^*(3\phi^*-2K_{yx})C_x^2 + \lambda^*\phi^*(\phi^*-K_{yx(2)})C_{x(2)}^2\right\}, \tag{3.11}$$

$$Var(T_0) = \bar{Y}^2\left\{\lambda C_y^2 + \lambda^* C_{y(2)}^2\right\}, \tag{3.12}$$

$$Var(T_1) = \bar{Y}^2\left[\lambda\left\{C_y^2 + (1-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + (1-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.13}$$

$$Var(T_2) = \bar{Y}^2\left[\lambda\left\{C_y^2 + (1-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2\right], \tag{3.14}$$

$$Var(T_3) = \bar{Y}^2\left[\lambda\left\{C_y^2 + 4(1-K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + (1-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.15}$$

$$Var(T_4) = \bar{Y}^2\left[\lambda\left\{C_y^2 + 4(1+K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + (1+2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.16}$$

$$Var(T_5) = \bar{Y}^2\left[\lambda\left\{C_y^2 + \phi'(\phi'-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + \phi'(\phi'-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.17}$$

$$Var(T_6) = \bar{Y}^2\left[\lambda\left\{C_y^2 + \phi'(\phi'-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2\right], \tag{3.18}$$

$$Var(T_7) = \bar{Y}^2\left[\lambda\left\{C_y^2 + 4\phi'(\phi'-K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + \phi'(\phi'-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.19}$$

$$Var(T_8) = \bar{Y}^2\left[\lambda\left\{C_y^2 + \phi^*(\phi^*-2K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + \phi^*(\phi^*-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.20}$$

$$Var(T_9) = \bar{Y}^2\left[\lambda\left\{C_y^2 + \phi^*(\phi^*-2K_{yx})C_x^2\right\} + \lambda^* C_{y(2)}^2\right], \tag{3.21}$$

$$Var(T_{10}) = \bar{Y}^2\left[\lambda\left\{C_y^2 + 4\phi^*(\phi^*-K_{yx})C_x^2\right\} + \lambda^*\left\{C_{y(2)}^2 + \phi^*(\phi^*-2K_{yx(2)})C_{x(2)}^2\right\}\right], \tag{3.22}$$

where $\phi' = \left(\dfrac{\bar{X}}{\bar{X}+C_x}\right)$ and $\phi^* = \left(\dfrac{\bar{X}}{\bar{X}+\rho}\right)$.

4. Efficiency comparison

The proposed class of estimators T is more efficient than

(i) usual unbiased estimator $T_0 = \bar{y}^*$ if

$$\left.\begin{aligned} 0<\alpha<\min.\left\{-\left(\frac{2K_{yx}}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\varphi}\right)\right\} \\ or\max.\left\{-\left(\frac{2K_{yx}}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\varphi}\right)\right\}<\alpha<0 \end{aligned}\right\}, \tag{4.1}$$

(ii) usual ratio estimator T_1 if

$$\left.\begin{aligned} 0<\alpha<\min.\left\{-\left(\frac{2K_{yx}-1}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}-1}{\varphi}\right)\right\} \\ or\ \max.\left\{-\left(\frac{2K_{yx}-1}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}-1}{\varphi}\right)\right\}<\alpha<0 \end{aligned}\right\}, \tag{4.2}$$

(iii) the ratio estimator T_2 if

$$\left.\begin{aligned} 0<\alpha<\min.\left\{-\left(\frac{2(K_{yx}-1)}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\varphi}\right)\right\} \\ or\quad \max.\left\{-\left(\frac{2(K_{yx}-1)}{\varphi}+\beta\right);-\left(\frac{2K_{yx(2)}}{\varphi}\right)\right\}<\alpha<0 \end{aligned}\right\}, \tag{4.3}$$

(iv) the ratio estimator T_3 if

$$\left.\begin{aligned} 0<\alpha<\min.\left[-\left\{\frac{2(K_{yx}-1)}{\varphi}+\beta\right\};-\left(\frac{2K_{yx(2)}-1}{\varphi}\right)\right] \\ or\quad \max.\left[-\left\{\frac{2(K_{yx}-1)}{\varphi}+\beta\right\};-\left(\frac{2K_{yx(2)}-1}{\varphi}\right)\right]<\alpha<0 \end{aligned}\right\}, \tag{4.4}$$

(v) the product estimator T_4 if

$$\left.\begin{aligned} 0<\alpha<\min.\left[-\left\{\frac{2(K_{yx}+1)}{\varphi}+\beta\right\};-\left(\frac{2K_{yx(2)}+1}{\varphi}\right)\right] \\ or\quad \max.\left[-\left\{\frac{2(K_{yx}+1)}{\varphi}+\beta\right\};-\left(\frac{2K_{yx(2)}+1}{\varphi}\right)\right]<\alpha<0 \end{aligned}\right\}, \tag{4.5}$$

(vi) the estimator T_5 if

$$\left.\begin{aligned} 0<\alpha<\min.\left[-\left\{\left(\frac{2K_{yx}-\varphi'}{\varphi}\right)+\beta\right\};-\left(\frac{2K_{yx(2)}-\varphi'}{\varphi}\right)\right] \\ or\quad \max.\left[-\left\{\left(\frac{2K_{yx}-\varphi'}{\varphi}\right)+\beta\right\};-\left(\frac{2K_{yx(2)}-\varphi'}{\varphi}\right)\right]<\alpha<0 \end{aligned}\right\}, \tag{4.6}$$

(vii) the estimator T_6 if

$$\left.\begin{array}{l} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\varphi'}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\varphi}\right)\right\} \\ or \quad \max.\left\{-\left(\dfrac{2K_{yx}-\varphi'}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\varphi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{4.7}$$

(viii) the estimator T_7 if

$$\left.\begin{array}{l} 0<\alpha<\min.\left[-\left\{\dfrac{2\left(K_{yx}-\varphi'\right)}{\varphi}+\beta\right\};\left(\dfrac{2K_{yx(2)}-\varphi'}{\varphi}\right)\right] \\ or \quad \max.\left[-\left\{\dfrac{2\left(K_{yx}-\varphi'\right)}{\varphi}+\beta\right\};\left(\dfrac{2K_{yx(2)}-\varphi'}{\varphi}\right)\right]<\alpha<0 \end{array}\right\}, \tag{4.8}$$

(ix) the estimator T_8 if

$$\left.\begin{array}{l} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\varphi^*}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}-\varphi^*}{\varphi}\right)\right\} \\ or \quad \max.\left\{-\left(\dfrac{2K_{yx}-\varphi^*}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}-\varphi^*}{\varphi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{4.9}$$

(x) the estimator T_9 if

$$\left.\begin{array}{l} 0<\alpha<\min.\left\{-\left(\dfrac{2K_{yx}-\varphi^*}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\varphi}\right)\right\} \\ or \quad \min.\left\{-\left(\dfrac{2K_{yx}-\varphi^*}{\varphi}+\beta\right);-\left(\dfrac{2K_{yx(2)}}{\varphi}\right)\right\}<\alpha<0 \end{array}\right\}, \tag{4.10}$$

(xi) the estimator T_{10} if

$$\left.\begin{array}{l} 0<\alpha<\min.\left[-\left\{\dfrac{2\left(K_{yx}-\varphi^*\right)}{\varphi}+\beta\right\};-\left(\dfrac{2K_{yx(2)}-\varphi^*}{\varphi}\right)\right] \\ or \quad \max.\left[-\left\{\dfrac{2\left(K_{yx}-\varphi^*\right)}{\varphi}+\beta\right\};-\left(\dfrac{2K_{yx(2)}-\varphi^*}{\varphi}\right)\right]<\alpha<0 \end{array}\right\}. \tag{4.11}$$

5. Empirical study

To look closely the excellence of the suggested estimators we consider two data sets:

Population I:

Source: Khare and Srivastava (1995)

The population of 100 consecutive trips (after leaving 20 outlier values) measured by two fuel meters for a small family car in normal usage given by Lewis et al (1991) has been taken into consideration. The measurements of turbine meter (in ml) are considered as main (study) character y and the measurements of displacement meter (in cm^3) are considered as auxiliary character x. We treat last 25% values as non-response values. The values of the parameters are as follows:

$\bar{Y} = 3500.12$, $\bar{X} = 260.84$, $C_y = 0.5941$, $C_x = 0.5996$, $\bar{Y}_2 = 3401.08$, $\bar{X}_2 = 259.96$, $C_{y(2)} = 0.4931$, $C_{x(2)} = 0.5151$, $\rho_{yx} = 0.985$, $\rho_{yx(2)} = 0.995$, $R = 13.4186$, $\beta_{yx} = 13.0961$, $\beta_{yx(2)} = 12.7820$, $W_2 = 0.25$, $N = 100$, $n = 25$

Population II:

Source: Khare and Srivastava (1993).

A list of 70 villages in India alongwith their populations in 1981 and cultivated areas (in acres) in the same year is considered; see Singh and Choudhary (1986). Here the cultivated areas (in acres) is taken as the main study variable and the population of the village is taken as the auxiliary variable. The parameters of the population are as follows:

$\bar{Y} = 981.29$, $\bar{X} = 1755.53$, $C_y = 0.6254$, $C_x = 0.8009$, $\bar{Y}_2 = 597.29$, $\bar{X}_2 = 1100.24$, $C_{y(2)} = 0.2488$, $C_{x(2)} = 0.3597$, $\rho_{yx} = 0.778$, $\rho_{yx(2)} = 0.445$, $R = 0.558971$, $\beta_{yx} = 0.339533$, $\beta_{yx(2)} = 0.172015$, $W_2 = 0.20$, $N = 70$, $n = 30$

Here, we have computed the Percent Relative Efficiencies (PREs) of the different suggested class of estimators with respect to the usual unbiased estimator $\bar{y}^*$ for different values of k.

Table 1: Percent Relative efficiency (P.RE (.)) of the estimators with respect to $\bar{y}^*$

	Estimators	$(1/k)$			
		$(1/5)$	$(1/4)$	$(1/3)$	$(1/2)$
Population I	T_0	100.00	100.00	100.00	100.00
	T_1	4791.40	4513.40	4185.87	3794.27
	T_2	188.12	216.44	271.55	425.73
	T_3	187.39	163.65	139.77	115.75
	T_4	15.27	14.51	13.61	12.49
	T_5	4800.47	4522.65	4195.21	3803.57
	T_6	188.14	216.46	271.59	425.85
	T_7	189.02	165.08	141.00	116.77
	T_8	4774.59	4499.75	4175.61	3787.56
	T_9	183.42	211.07	264.94	415.86
	T_{10}	186.25	162.65	138.92	115.04
	T_{opt}	5139.86	4794.01	4395.16	3930.13
Population II	T_0	100.00	100.00	100.00	100.00
	T_1	123.1481	129.00	135.87	144.05
	T_2	147.78	151.76	156.25	161.33
	T_3	29.49	29.12	28.74	28.34
	T_4	9.27	9.12	8.96	8.80
	T_5	114.66	120.56	127.55	135.96
	T_6	133.00	135.79	138.89	142.35
	T_7	29.23	28.92	28.60	28.27
	T_8	114.66	120.56	127.55	135.96
	T_9	133.00	135.79	138.89	142.35
	T_{10}	29.22	28.92	28.59	28.26
	T_{opt}	255.77	263.27	271.74	281.39

From table 1, we apprehend that the proposed class of estimators under optimum condition is more desirable over all the considered estimators in two populations I and II.

Population I: It is observed from the table 1 that the percent relative efficiencies of the estimators T_1, T_3, T_4, T_5, T_7, T_8, T_{10} and T_{opt} decreases as $(1/k)$ increases, but for the estimators T_2, T_6 and T_9, it increases with the increase in the value of $(1/k)$. The performance of the estimator T_4 is worse than that of the usual unbiased estimator $\bar{y}^*$.

Population II: Table 1 envisaged that the PREs of the estimators T_1, T_2, T_5, T_6, T_8, T_9 and T_{opt} increases as the value of $(1/k)$ increases, whereas PREs of the estimators T_3, T_4, T_7 and T_{10} decreases as $(1/k)$ increases. Among the estimators T_0, T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9 and T_{10}, the proposed class of estimators under optimum condition (T_{opt}) is the best in the sense of having the smallest variance. Overall, estimator T_{opt} has the smallest gain in efficiency. The performance of the estimators T_3, T_4, T_7 and T_{10} are worse than that of the usual unbiased estimator $\bar{y}^*$.

Thus the class of estimator T at its optimum value is recommended for further use in practice.

Acknowledgement

Author is thankful to the learned referees for their valuable constructive suggestions regarding improvement of the paper.

References

1. Cochran, W.G. (1977).*Sampling Techniques, 3rd ed..* John Wiley and Sons, New York.
2. El-Badry, M.A. (1956). A Sampling procedure for mailed Questionnaires.J. *Amer. Statist. Assoc.,* 51, 209-227.
3. Hansen, M.H., and Hurwitz, W.N. (1946). The problem of non-response in samplesurveys. *J. Amer. Statist. Assoc.,*41,517- 529.
4. Khare, B.B., and Srivastava, S. (1993). Estimation of population mean using auxiliarycharacter in presence of non-response. *Nat. Acad. Sc. Letters, India*, 16(3), 111-114.
5. Khare, B.B., and Srivastava, S. (1995). Study of conventional and alternative two-phasesampling ratio, product and regression estimators in presence of non-response. *Proc. Nat. Acad. Sci. India,*65(A), II, 195-203.
6. Khare, B.B., and Srivastava, S. (1997). Transformed ratio type estimators for thepopulation mean in the presence of non-response. *Comm. Statist. - Theory Methods*, 26(7), 1779-1791.

7. Lewis, P.A., Jones, P.W., Polak, J.W., and Tillotson, H.T. (1991). The problem ofconversion in method comparison studies. *Applied Statistics*, 40(1), 105-112.

8. Okafor, F.C., and Lee, H. (2000). Double sampling for ratio and regression estimationwith sub sampling the non-respondent. *Survey Methodology*, 26, 183-188.

9. Rao, P.S.R.S. (1986). Ratio Estimation with sub sampling the non-respondents. *Survey Methodology,* 12(2), 217-230.

10. Rao, P. S. R. S. (1987). Ratio and regression estimates with sub-sampling the non-respondents. *Paper presented at a special contributed session of the International Statistical Association Meetings, September, 2-16, Tokyo, Japan.*

11. Singh, H.P. and Kumar, S. (2008a).A regression approach to the estimation of finitepopulation mean in presence of non-response. *Aust. N. Z. J. Stat.,* 50(4)*, 395-408*.

12. Singh, H.P., and Kumar, S. (2008b). A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in the presence of non-response. *J. of Statistical Theory and Practice,* 2(4)*, 677-692.*

13. Singh, H.P., and Kumar, S. (2009). A general class of estimators of the population mean in survey sampling using auxiliary information with sub sampling the non - respondents. *The Korean J. of Applied Statistics*, 22(2), 387- 402.

14. Singh, H.P., and Kumar, S. (2010a). Improved estimation of population mean under two phase sampling with sub sampling the non-respondents. J. S. P. I., 140, 9, 2536-2550.

15. Singh, H.P., and Kumar, S. (2010 b). Estimation of mean in presence of non-responseusing two phase sampling scheme. *Statistical Papers*, 50(3), 559-582.

16. Singh, H.P., and Kumar, S. (2011). Combination of ratio and regression estimators in presence of non-response. Brazilian Journal of Probability and Statistics, 25(2), 205-217.

17. Sodipo, A.A., and Obisesan, K.O. (2007). Estimation of the population mean usingdifference cum ratio estimator with full response on the auxiliary character. *Res. J. Applied Sci*., 2(6), 769-772.

18. Tabasum, R. and Khan, I.A. (2004). Double sampling for ratio estimation with nonresponse. *J. Ind. Soc. Agril. Statist.* 58(3), 300-306.

19. Tabasum, R. and Khan, I.A. (2006).Double sampling ratio estimator for the population mean in presence of non-response. *Assam Statist. Review,* 20(1), 73-83.