STATISTICS IN TRANSITION new series, June 2018
Vol. 19, No. 2, pp. 259–276, DOI 10.21307/stattrans-2018-015

A GENERALIZED EXPONENTIAL TYPE ESTIMATOR OF POPULATION MEAN IN THE PRESENCE OF NON-RESPONSE

Siraj Muneer[1], Javid Shabbir[2], Alamgir Khalil[3]

ABSTRACT

In this article, we propose a class of generalized exponential type estimators to estimate the finite population mean by using two auxiliary variables under non-response in simple random sampling. The proposed estimator under non-response in different situations has been studied and gives minimum mean square error as compared to all other considered estimators. Usual exponential ratio type estimator, exponential product type estimator and many more estimators are also identified from the proposed estimator. We use three real data sets to obtain the efficiencies of estimators.

Key words: auxiliary variables, bias, MSE, efficiency, non-response.

1. Introduction

In survey sampling, it is assumed that all the observations are correctly measured on the characteristics under study. But in practice, when we fail to collect the complete information on different variables, non-response is supposed to occur. Non-response occurs due to many reasons, which includes the lack of information provided by respondents, also some of the respondents refuse to answer the questionnaire, sometimes it is difficult to find out the respondents, etc. The common approach to overcome non-response problem is to contact the non-respondents and obtain maximum information as much as possible. Generally auxiliary information is used to increase the precision of the estimators when there exists a correlation between the study and the auxiliary variables. Ratio, product and regression estimators are good examples in this context. In daily life there are many situations when we are unable to access the complete information either on the study variable or the auxiliary variable or at the same time both on the study and the auxiliary variables. Hansen and Hurwitz (1946) were the first to suggest a non-response handling technique in mail surveys combined the advantages of mailed questionnaires and personal interviews. Later on, several authors, including Srinath (1971), selected the subsample of non-respondents, where the sub-sampling

[1]Department of Statistics, University of Peshawar, Peshawar, Pakistan.
E-mail: sirajmuneer1@gmail.com
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
E-mail: javidshabbir@gmail.com
[3]Department of Statistics, University of Peshawar, Peshawar, Pakistan.
E-mail: profalamgir@yahoo.com

fraction varied according to the non-response rates. Rao (1986) suggested a ratio type estimator for population mean, in the presence of non-response for two-phase sampling, when the population mean of the auxiliary variable x was unknown and non-response occurred on the auxiliary variable. Khare and Srivastava (1993, 1995, 1997, 2010) and Olkin (1958) suggested different ratio and product types estimators for estimation of population mean using the auxiliary information under non-response. Similarly, Singh and Kumar (2008) and Singh et al. (2010) have made significant contributions and proposed ratio, product and difference classes of estimators under non-response. El-Badry (1956), Bahl and Tuteja (1991), Kumar and Bhougal (2011), Muneer et al. (2017) and Ismail et al. (2011) suggested many estimators in two-phase sampling with sub-sampling of non-respondents in estimating the finite population mean. Khare and Sinha (2007, 2009, 2011) proposed some classes of estimators for estimating population mean in the presence of non-response using multi-auxiliary characters in different ways. For controlling the non-response bias and eliminating the need for call backs in survey sampling, Tabasum and Khan (2006), Shabbir and Nasir (2013) and references cited therein have discussed some good techniques and plans for the estimation of finite population mean followed the technique proposed by Hansen and Hurwitz (1946) using one or more auxiliary variables in the presence of non-response. Now, we explain the Hansen and Hurwitz (1946) strategy for non-response.
Suppose a finite population $U=(1,2,...,N)$ of size N units can be divided into two classes $N=N_1+N_2$. Let N_1 and N_2 be the number of units in the population that form the response and non-response classes respectively. We draw a sample of size n units from U by using a simple random sample without replacement (SRSWOR) sampling scheme. Let r_1 units respond and $r_2=(n-r_1)$ units do not respond in the first attempt. Let a subsample of size k_2 units be selected from r_2 non-respondents units, such that $r_2=k_2h,(h>1)$. Let y_i and (x_i,z_i) be the values of the study variable (y) and the auxiliary variables (x,z) respectively. Let $\bar{y}$ and $(\bar{x},\bar{z})$ be the sample means corresponding to population means $\bar{Y}$ and $(\bar{X},\bar{Z})$ respectively.

2. Notations and Symbols with Selected Estimators

To obtain the properties of estimators, we define the following symbols and notations.
Let $e_0^*=\left(\frac{\bar{y}^*-\bar{Y}}{\bar{Y}}\right), e_1^*=\left(\frac{\bar{x}^*-\bar{X}}{\bar{X}}\right), e_1=\left(\frac{\bar{x}-\bar{X}}{\bar{X}}\right)$, $e_2=\left(\frac{\bar{z}-\bar{Z}}{\bar{Z}}\right)$, $e_2^*=\left(\frac{\bar{z}^*-\bar{Z}}{\bar{Z}}\right)$
are the relative error terms, such that
$E(e_i{}^*)=0,(i=0,1,2)$, $E(e_i)=0,(i=1,2)$.
$E(e_0^{*2})=\lambda C_y^2+\theta C_{y(2)}^2=V_y^*$, $\quad E(e_1^{*2})=\lambda C_x^2+\theta C_{x(2)}^2=V_x^*$,
$E(e_2^{*2})=\lambda C_z^2+\theta C_{z(2)}^2=V_z^*$, $\quad E(e_0^*e_1^*)=\lambda C_{yx}+\theta C_{yx(2)}=V_{yx}^*$,
$E(e_0^*e_2^*)=\lambda C_{yz}+\theta C_{yz(2)}=V_{yz}^*, E(e_1^*e_2^*)=\lambda C_{xz}+\theta C_{xz(2)}=V_{xz}^*$,
$E(e_1^*e_1)=E(e_1^2)=\lambda C_x^2=V_x, E(e_0^*e_1)=E(e_0e_1)=\lambda C_{yx}=V_{yx}$,
where

$C_y^2 = \frac{S_y^2}{\bar{Y}^2}$, $C_x^2 = \frac{S_x^2}{\bar{X}^2}$, $C_z^2 = \frac{S_z^2}{\bar{Z}^2}$, $C_{yx} = \rho_{yx} C_y C_x$, $C_{yz} = \rho_{yz} C_y C_z$,
$C_{xz} = \rho_{xz} C_x C_z$, $C_{y(2)}^2 = \frac{S_{y(2)}^2}{\bar{Y}^2}$, $C_{x(2)}^2 = \frac{S_{x(2)}^2}{\bar{X}^2}$, $C_{z(2)}^2 = \frac{S_{z(2)}^2}{\bar{Z}^2}$,
$C_{yx(2)} = \rho_{yx(2)} C_{y(2)} C_{x(2)}$, $C_{yz(2)} = \rho_{yz(2)} C_{y(2)} C_{z(2)}$, $C_{xz(2)} = \rho_{xz(2)} C_{x(2)} C_{z(2)}$,
$\lambda = (\frac{1-f}{n})$, $f = \frac{n}{N}$, $\theta = W_2(\frac{h-1}{n})$ and $W_2 = \frac{N_2}{N}$.

Now, we review some important estimators which are available in the literature.

1. Hansen and Hurwitz (1946) were the first who formulated an unbiased estimator of the population mean $\bar{Y}$ of the study variable Y in the presence of non-response. Initially they considered the mailed survey in the first attempt and personal interviews in the second attempt after the deadline was over. The estimator is given by

$$\bar{y}^* = \left(\frac{r_1}{n}\right)\bar{y}_{r_1} + \left(\frac{r_2}{n}\right)\bar{y}_{k_2}, \tag{1}$$

where $\bar{y}_{r_1} = \frac{1}{r_1}\sum_{i=1}^{r_1} y_i$ and $\bar{y}_{k_2} = \frac{1}{k_2}\sum_{i=1}^{k_2} y_i$.
The variance of $\bar{y}^*$, is given by

$$V(\bar{y}^*) = \bar{Y}^2 V_y^*. \tag{2}$$

2. The ratio and product estimators under non-response case.
When non-response exists on the study variable y as well as on the auxiliary variable x, the traditional ratio and product estimators for population mean $\bar{Y}$ are given by

$$\bar{y}_R^* = \bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right), \tag{3}$$

and

$$\bar{y}_P^* = \bar{y}^*\left(\frac{\bar{x}^*}{\bar{X}}\right), \tag{4}$$

where $\bar{y}^*$ and $\bar{x}^*$ are the Hansen and Hurwitz (1946) estimators for population means $\bar{Y}$ and $\bar{X}$ respectively and are defined by $\bar{y}^* = \left(\frac{r_1}{n}\right)\bar{y}_{r_1} + \left(\frac{r_2}{n}\right)\bar{y}_{k_2}$ and $\bar{x}^* = \left(\frac{r_1}{n}\right)\bar{x}_{r_1} + \left(\frac{r_2}{n}\right)\bar{x}_{k_2}$ with $(\bar{y}_{r_1}, \bar{x}_{r_1})$ and $(\bar{y}_{k_2}, \bar{x}_{k_2})$ are the sample means of (y, x) based on the samples of r_1 and k_2 units respectively.

The $MSEs$ of $\bar{y}_R^*$ and $\bar{y}_P^*$, to the first order of approximation, are given by

$$MSE(\bar{y}_R^*) \cong \bar{Y}^2(V_y^* + V_x^* - 2V_{yx}^*), \tag{5}$$

and

$$MSE(\bar{y}_P^*) \cong \bar{Y}^2(V_y^* + V_x^* + 2V_{yx}^*). \tag{6}$$

We observed that $\bar{y}_R^*$ and $\bar{y}_P^*$ perform better than $\bar{y}^*$, if $V_{yx}^* > \frac{1}{2}V_x^*$ and $V_{yx}^* < -\frac{1}{2}V_x^*$ respectively.

3. Rao (1986) suggested ratio and product estimators under no-response. When non-response exists only on the study variable y, while the complete information on the auxiliary variable x is available, the ratio and product estimators, are given by

$$\bar{y}^*_{Rao(R)} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right), \tag{7}$$

and

$$\bar{y}^*_{Rao(P)} = \bar{y}^* \left(\frac{\bar{x}}{\bar{X}}\right), \tag{8}$$

where $\bar{x}$ is the sample mean $\bar{X}$ based on complete information, and $\bar{X}$ is the population mean of the auxiliary variable.

The $MSEs$ of $\bar{y}^*_{Rao(R)}$ and $\bar{y}^*_{Rao(P)}$, to the first order of approximation, are given by

$$MSE(\bar{y}^*_{Rao(R)}) \cong \bar{Y}^2(V^*_y + V_x - 2V_{yx}), \tag{9}$$

and

$$MSE(\bar{y}^*_{Rao(P)}) \cong \bar{Y}^2(V^*_y + V_x + 2V_{yx}). \tag{10}$$

Note that $\bar{y}^*_{Rao(R)}$ and $\bar{y}^*_{Rao(P)}$ perform better then $\bar{y}^*$ if $V_{yx} > \frac{1}{2}V_x$ and $V_{yx} < -\frac{1}{2}V_x$ respectively.

4. Bahl and Tuteja (1991) exponential ratio and product type estimators for population mean $\bar{Y}$, when non-response exists on the study variable y as well as on the auxiliary variable x as:

$$\bar{y}^*_{exp(R)} = \bar{y}^* \exp\left(\frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*}\right), \tag{11}$$

and

$$\bar{y}^*_{exp(P)} = \bar{y}^* \exp\left(\frac{\bar{x}^* - \bar{X}}{\bar{x}^* + \bar{X}}\right). \tag{12}$$

The $MSEs$ of $\bar{y}^*_{exp(R)}$ and $\bar{y}^*_{exp(P)}$, to the first order of approximation, are given by

$$MSE(\bar{y}^*_{exp(R)}) \cong \bar{Y}^2\left(V^*_y + \frac{1}{4}V^*_x - V^*_{yx}\right), \tag{13}$$

and

$$MSE(\bar{y}^*_{exp(P)}) \cong \bar{Y}^2\left(V^*_y + \frac{1}{4}V^*_x + V^*_{yx}\right). \tag{14}$$

Both the estimators $\bar{y}^*_{exp(R)}$ and $\bar{y}^*_{exp(P)}$ are more efficient than $\bar{y}^*$ if $V^*_{yx} > \frac{1}{4}V^*_x$ and $V^*_{yx} < -\frac{1}{4}V^*_x$ respectively.

5. Singh and Kumar (2008) suggested ratio, product and difference type estima-

tors in the case of non-response. They considered the situation in which the population mean of the auxiliary variable x is known, but some units fail to provide information on the study variable y and the auxiliary variable x. The estimator is given by:

$$\bar{y}^*_{SK(R1)} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)\left(\frac{\bar{X}}{\bar{x}}\right), \tag{15}$$

where $\bar{x}^*$ and $\bar{x}$, both are unbiased estimators of the population mean $\bar{X}$ of the auxiliary variable x.
The *MSE* of $\bar{y}^*_{SK(R1)}$, to the first order of approximation, is given by

$$MSE(\bar{y}^*_{SK(R1)}) \cong \bar{Y}^2 \left(V^*_y + V^*_x + 3V_x - 2(V^*_{yx} + V_{yx})\right). \tag{16}$$

Note that $\bar{y}^*_{SK(R1)}$ performs better than $\bar{y}^*$ if $(V^*_{yx} + V_{yx}) > \frac{1}{2}(V^*_x + V_x)$. The product estimator of the above mentioned situation is

$$\bar{y}^*_{SK(P)} = \bar{y}^* \left(\frac{\bar{x}^*}{\bar{X}}\right)\left(\frac{\bar{x}}{\bar{X}}\right). \tag{17}$$

The *MSE* of $\bar{y}^*_{SK(P)}$, to the first order of approximation, is given by

$$MSE(\bar{y}_{SK(P)}) \cong \bar{Y}^2 \left(V^*_y + V^*_x + 3V_x + 2(V^*_{yx} + V_{yx})\right). \tag{18}$$

Note that *MSE* of $\bar{y}^*_{SK(P)}$, is smaller than $\bar{y}^*$ if $V^*_{yx} + V_{yx} < -\frac{1}{2}(V^*_x + 3V_x)$.
Singh and Kumar (2008) also suggested the generalized ratio-type estimator of the above mentioned situations as

$$\bar{y}^*_{SK(R2)} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)^{\alpha_1}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha_2}, \tag{19}$$

where α_1 and α_2 are constants whose values are to be determined.

The minimum *MSE* of $\bar{y}^*_{SK(R2)}$ to the first order of approximation, at optimum values of α_1 and α_2 *i.e.* $\alpha_{1opt} = \frac{(V^*_{yx} - V_{yx})}{(V^*_x - V_x)}$ and $\alpha_{2opt} = \frac{(V^*_x V_{yx} - V^*_{yx} V_x)}{V_x(V^*_x - V_x)}$, is given by

$$MSE(\bar{y}^*_{SK(R2)})_{min} \cong \bar{Y}^2 \left[V^*_y - \frac{V_x V^{*2}_{yx} + V^*_x V^2_{yx} - 2V_{yx} V^*_{yx} V_x}{V_x(V^*_x - V_x)}\right]. \tag{20}$$

Note that $\bar{y}^*_{SK(R2)}$ performs better than $\bar{y}^*$
if $\frac{V_x V^{*2}_{yx} + V^*_x V^2_{yx} - 2V_{yx} V^*_{yx} V_x}{V_x(V^*_x - V_x)} > 0$
Singh and Kumar (2008) also suggested a difference type estimator in the case of non-response as

$$\bar{y}^*_{SK(d)} = \bar{y}^* + d_1(\bar{x} - \bar{x}^*) + d_2(\bar{X} - \bar{x}), \tag{21}$$

where d_1 and d_2 are constants whose values are to be determined.
The minimum *MSE* of $\bar{y}^*_{SK(d)}$ to the first order of approximation, at optimum values of d_1 and d_2 *i.e.* $d_{1opt} = \frac{\bar{Y}(V^*_{yx}-V_{yx})}{\bar{X}(V^*_x-V_x)}$ and $d_{2opt} = \frac{\bar{Y}V_{yx}}{\bar{X}V_x}$, is given by,

$$MSE(\bar{y}^*_{SK(d)})_{min} \cong \bar{Y}^2 \left[V^*_y - \frac{(V^*_{yx}-V_{yx})^2}{(V^*_x - V_x)} - \frac{V^2_{yx}}{V_x} \right]. \tag{22}$$

Note that $\bar{y}^*_{SK(d)}$ performs better than $\bar{y}^*$ if $\left[\frac{(V^*_{yx}-V_{yx})^2}{(V^*_x-V_x)} + \frac{V^2_{yx}}{V_x} \right] > 0$. Also, we observed that $MSE(\bar{y}^*_{SK(d)})_{min} = MSE(\bar{y}^*_{SK(R2)})_{min}$.

6. Kumar and Bhougal (2011) proposed ratio-product-type exponential estimator for the population mean $\bar{Y}$, when non-response exists on both the study variable y and the auxiliary variable x as

$$\bar{y}^*_{KB} = \bar{y}^* \left[\alpha \exp\left(\frac{\bar{X} - \bar{x}^*}{\bar{X} + \bar{x}^*} \right) + (1-\alpha) \exp\left(\frac{\bar{x}^* - \bar{X}}{\bar{x}^* + \bar{X}} \right) \right], \tag{23}$$

where α is a constant whose value is to be determined.
The minimum *MSE* of $\bar{y}^*_{KB}$ at optimum value of $\alpha_{opt} = \frac{1}{2}(1 + 2\frac{V^*_{yx}}{V^*_x})$, to the first order of approximation, is given by

$$MSE(\bar{y}^*_{KB})_{min} \cong \bar{Y}^2 \left(V^*_y - \frac{V^{*2}_{yx}}{V^*_x} \right). \tag{24}$$

Note that $\bar{y}^*_{KB}$ performs better than $\bar{y}^*$ if $\frac{V^{*2}_{yx}}{V^*_x} > 0$, which is always true.

3. Class of Estimators

In application, our purpose was to construct a type of general class of estimators which contains many estimators, stable and efficient. So motivated by Singh and Shukla (1993) and Shukla et al. (2012), we propose the following general class of estimators in the case of non-response exists on the study variable as well as on the two auxiliary variables. Initially Bahl and Tuteja (1991) gave the idea of exponential ratio type and product type estimators for estimating the population mean by using the single auxiliary variable. Also, we can generate many more estimators by substituting different values of $(K_i, i = 1,2,3,4)$. The proposed estimator is constructed by combining the ideas of Bahl and Tuteja (1991), Singh and Shukla (1993) and Shukla (2012), given by

$$\bar{y}^*_{prop} = \bar{y}^* \left[\exp\left(\frac{G_1 - D_1}{G_1 + D_1} \right) \exp\left(\frac{G_2 - D_2}{G_2 + D_2} \right) \right], \tag{25}$$

where "*prop*" indicates proposed. $G_1 = (A_1 + C_1)\bar{X} + fB_1\bar{x}^*$, $G_2 = (A_2 + C_2)\bar{Z} + fB_2\bar{z}^*$,

$D_1=(A_1+fB_1)\bar{X}+C_1\bar{x}^*$, $D_2=(A_2+fB_2)\bar{Z}+C_2\bar{z}^*$, $A_i=(K_i-1)(K_i-2)$, $B_i=(K_i-1)(K_i-4)$, $C_i=(K_i-2)(K_i-3)(K_i-4)$, $(i=1,2,3,4)$.
Substituting different values of K_i in (25), we can generate many more different types of estimators from our proposed class of estimators given in Table 1.

Table 1: Some members of the proposed class of family of estimators $\bar{y}^*_{prop}$

Estimators	Estimators
$K_1=1$ and $K_2=1$ $\bar{y}^*_{prop1}=\bar{y}^*\exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right)\exp\left(\frac{\bar{Z}-\bar{z}^*}{\bar{Z}+\bar{z}^*}\right)$	$K_1=1$ and $K_2=2$ $\bar{y}_{prop2}=\bar{y}^*\exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right)\exp\left(\frac{\bar{z}^*-\bar{Z}}{\bar{z}^*+\bar{Z}}\right)$
$K_1=1$ and $K_2=3$ $\bar{y}^*_{prop3}=\bar{y}^*\exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right)\exp\left(\frac{n(\bar{Z}-\bar{z}^*)}{2N\bar{Z}-n(\bar{z}^*+\bar{Z}})\right)$	$K_1=1$ and $K_2=4$ $\bar{y}^*_{prop4}=\bar{y}^*\exp\left(\frac{\bar{X}-\bar{x}^*}{\bar{X}+\bar{x}^*}\right)$
$K_1=2$ and $K_2=1$ $\bar{y}^*_{prop5}=\bar{y}^*\exp\left(\frac{\bar{x}^*-\bar{X}}{\bar{x}^*+\bar{X}}\right)\exp\left(\frac{\bar{Z}-\bar{z}^*}{\bar{Z}+\bar{z}^*}\right)$	$K_1=2$ and $K_2=2$ $\bar{y}^*_{prop6}=\bar{y}^*\exp\left(\frac{\bar{x}^*-\bar{X}}{\bar{x}^*+\bar{X}}\right)\exp\left(\frac{\bar{z}^*-\bar{Z}}{\bar{z}^*+\bar{Z}}\right)$
$K_1=2$ and $K_2=3$ $\bar{y}^*_{prop7}=\bar{y}^*\exp\left(\frac{\bar{x}^*-\bar{X}}{\bar{x}^*+\bar{X}}\right)\exp\left(\frac{n(\bar{Z}-\bar{z}^*)}{2N\bar{Z}-n(\bar{z}^*+\bar{Z}})\right)$	$K_1=2$ and $K_2=4$ $\bar{y}^*_{prop8}=\bar{y}^*\exp\left(\frac{\bar{x}^*-\bar{X}}{\bar{x}^*+\bar{X}}\right)$
$K_1=3$ and $K_2=1$ $\bar{y}^*_{prop9}=\bar{y}^*\exp\left(\frac{n(\bar{X}-\bar{x}^*)}{2N\bar{X}-n(\bar{x}^*+\bar{X})}\right)\exp\left(\frac{\bar{Z}-\bar{z}^*}{\bar{Z}+\bar{z}^*}\right)$	$K_1=3$ and $K_2=2$ $\bar{y}^*_{prop10}=\bar{y}^*\exp\left(\frac{n(\bar{X}-\bar{x}^*)}{2N\bar{X}-n(\bar{x}^*+\bar{X})}\right)\exp\left(\frac{\bar{z}^*-\bar{Z}}{\bar{z}^*+\bar{Z}}\right)$
$K_1=3$ and $K_2=3$ $\bar{y}^*_{prop11}=\bar{y}^*\exp\left(\frac{n(\bar{X}-\bar{x}^*)}{2N\bar{X}-n(\bar{x}^*+\bar{X})})\right)\exp\left(\frac{n(\bar{Z}-\bar{z}^*)}{2N\bar{Z}-n(\bar{z}^*+\bar{Z}})\right)$	$K_1=3$ and $K_2=4$ $\bar{y}^*_{prop12}=\bar{y}^*\exp\left(\frac{n(\bar{X}-\bar{x}^*)}{2N\bar{X}-n(\bar{x}^*+\bar{X})}\right)$
$K_1=4$ and $K_2=1$ $\bar{y}^*_{prop13}=\bar{y}^*\exp\left(\frac{\bar{Z}-\bar{z}^*}{\bar{Z}+\bar{z}^*}\right)$	$K_1=4$ and $K_2=2$ $\bar{y}^*_{prop14}=\bar{y}^*\exp\left(\frac{\bar{z}^*-\bar{Z}}{\bar{z}^*+\bar{Z}}\right)$
$K_1=4$ and $K_2=3$ $\bar{y}^*_{prop15}=\bar{y}^*\exp\left(\frac{n(\bar{Z}-\bar{z}^*)}{2N\bar{Z}-n(\bar{z}^*+\bar{Z}})\right)$	$K_1=4$ and $K_2=4$ $\bar{y}^*_{prop16}=\bar{y}^*$

Solving $\bar{y}^*_{prop}$ given in Eq. (25) in terms of $e's$ (defined earlier in Section 2), we have

$$\bar{y}^*_{prop} \cong \bar{Y}(1+e_0^*)\left(1+\frac{1}{2}\sigma_1 e_1^*-\frac{1}{4}\sigma_1\nu_1 e_1^{*2}+\frac{1}{8}\sigma_1^2 e_1^{*2}+....\right)$$
$$\left(1+\frac{1}{2}\sigma_2 e_2^*-\frac{1}{4}\sigma_2\nu_2 e_2^{*2}+\frac{1}{8}\sigma_2^2 e_2^{*2}+....\right), \tag{26}$$

where $\sigma_1=\frac{fB_1-C_1}{A_1+fB_1+C_1}$, $\nu_1=\frac{fB_1+C_1}{A_1+fB_1+C_1}$, $\sigma_2=\frac{fB_2-C_2}{A_2+fB_2+C_2}$, $\nu_2=\frac{fB_2+C_2}{A_2+fB_2+C_2}$.
To first order of approximation, we have

$$\bar{y}^*_{prop} - \bar{Y} \cong \bar{Y}(e^*_0 + \frac{1}{2}\sigma_1 e^*_1 + \frac{1}{2}\sigma_2 e^*_2 + \frac{1}{2}\sigma_1 e^*_0 e^*_1 + \frac{1}{2}\sigma_2 e^*_0 e^*_2 - \frac{1}{4}\sigma_1 \nu_1 e^{*2}_1 - \frac{1}{4}\sigma_2 \nu_2 e^{*2}_2 + \frac{1}{8}\sigma^2_1 e^{*2}_1 + \frac{1}{8}\sigma^2_2 e^{*2}_2). \quad (27)$$

The bias of $\bar{y}^*_{prop}$ to the first order of approximation, is given by

$$B(\bar{y}^*_{prop}) \cong \bar{Y}\left[\frac{1}{2}\sigma_1 V^*_{yx} + \frac{\sigma_1}{2}\sigma_2 V^*_{yz} + \frac{1}{8}V^*_x(\sigma^2_1 - 2\sigma_1\nu_1) + \frac{1}{8}V^*_z(\sigma^2_2 - 2\sigma_2\nu_2)\right]. \quad (28)$$

The *MSE* of $\bar{y}^*_{prop}$ to the first order of approximation, is given by

$$MSE(\bar{y}^*_{prop}) \cong \bar{Y}^2 E\left[e^*_0 + \frac{1}{2}\sigma_1 e^*_1 + \frac{1}{2}\sigma_2 e^*_2\right]^2.$$

Solving above equation, we have

$$MSE(\bar{y}^*_{prop}) \cong \bar{Y}^2\left[V^*_y + \frac{1}{4}\sigma^2_1 V^*_x + \frac{1}{4}\sigma^2_2 V^*_z + \sigma_1 V^*_{yx} + \sigma_2 V^*_{yz} + \frac{1}{2}\sigma_1\sigma_2 V^*_{xz}\right]. \quad (29)$$

Differentiate Eq.(29) with respect to σ_1 and σ_2, we get the optimum values of σ_1 and σ_2 i.e.

$$\sigma_{1(opt)} = \frac{2(V^*_{yz}V^*_{xz} - V^*_{yx}V^*_z)}{V^*_x V^*_z - V^{*2}_{xz}}$$

and

$$\sigma_{2(opt)} = \frac{2(V^*_{yx}V^*_{xz} - V^*_{yz}V^*_x)}{V^*_x V^*_z - V^{*2}_{xz}}.$$

Substituting the optimum values of $\sigma_{1(opt)}$ and $\sigma_{2(opt)}$ in Eq.(29), we get minimum *MSE* of $\bar{y}^*_{prop}$, given by

$$MSE(\bar{y}^*_{prop})_{min} \cong \bar{Y}^2\left[V^*_y - \frac{V^{*2}_{yx}V^*_z + V^{*2}_{yz}V^*_x - 2V^*_{yx}V^*_{yz}V^*_{xz}}{V^*_x V^*_z - V^{*2}_{xz}}\right]. \quad (30)$$

4. Theoretical Comparison

A comparison of our MSE estimator and previously presented 12 different estimators is given as

• By variance of Hansen and Hurwitz (1946) estimator and our MSE estimator:
$MSE(\bar{y}^*_{prop})_{min} < V(\bar{y}^*)$ if
$\frac{A_1}{A_2} > 0$, where
$A_1 = (V^{*2}_{yx}V^*_z + V^{*2}_{yz}V^*_x - 2V^*_{yx}V^*_{yz}V^*_{xz})$ and $A_2 = (V^*_x V^*_z - V^{*2}_{xz})$.

• By MSE of Rao (1986) estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{Rao(R)})$ if

$(V_x - 2V_{yx}) + \frac{A_1}{A_2} > 0.$

•By MSE of Rao (1986) estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{Rao(P)})$ if

$(V_x + 2V_{yx}) + \frac{A_1}{A_2} > 0.$

• By MSE of ratio estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_R)$ if

$(V^*_x - 2V^*_{yx}) + \frac{A_1}{A_2} > 0.$

• By MSE of product estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_P)$ if

$(V^*_x + 2V^*_{yx}) + \frac{A_1}{A_2} > 0.$

• By MSE of Bahl and Tuteja (1991) exponential ratio estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{exp(R)})$ if

$(\frac{1}{4}V^*_x - V^*_{yx}) + \frac{A_1}{A_2} > 0.$

• By MSE of Bahl and Tuteja (1991) exponential product estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{exp(P)})$ if

$(\frac{1}{4}V^*_x + V^*_{yx}) + \frac{A_1}{A_2} > 0.$

• By MSE of Singh and Kumar (2008) ratio type estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{SK(R1)})$ if

$[V^*_x + 3V_x - 2(V^*_{yx} + V_{yx})] + \frac{A_1}{A_2} > 0.$

• By MSE of Singh and Kumar (2008) product type estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{SK(P)})$ if

$[V^*_x + 3V_x + 2(V^*_{yx} + V_{yx})] + \frac{A_1}{A_2} > 0.$

• By MSE of Kumar and Bhougal (2011) ratio and product type estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{KB})_{min}$ if

$\frac{A_1}{A_2} - \frac{V^{*2}_{yx}}{V^*_x} > 0.$

• By MSE of Singh and Kumar (2008) chain ratio type estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{SK(R2)})_{min}$ if

$\frac{A_1}{A_2} - \frac{B_1}{B_2} > 0$, where

$B_1 = (V_{yx}^{*2} V_x + V_{yx}^2 V_x^* - 2V_{yx} V_{yx}^* V_x)$ and $B_2 = V_x(V_x^* - V_x)$.

• By MSE of Singh and Kumar (2008) difference type estimator and our MSE estimator:

$MSE(\bar{y}^*_{prop})_{min} < MSE(\bar{y}^*_{SK(d)})_{min}$ if

$$\frac{A_1}{A_2} - \frac{(V_{yx}^* - V_{yx})^2}{(V_x^* - V_x)} - \frac{V_{yx}^2}{V_x} > 0.$$

Note: The proposed class of estimators performs better than all other considered estimators, if the above mentioned conditions $(i) - (Xii)$ are satisfied.

5. Numerical Comparison

To observe the performance of our proposed generalized class of estimators with respect to other considered estimators, we use the following data sets, which were earlier used by many authors in the literature. We used different values of *h*, i.e. 2, 4, 6, 8 and 16, in Tables 2-4 in our study.

1. Data set 1 [Source: Khare and Sinha (2007)]

   y: Weights of the children in kilograms.
   x: Skull circumference of the children in centimeter.
   z: Chest circumference of the children in centimeter.

   $N = 95, n = 30, W_2 = 0.25, \bar{Y} = 19.4968, \bar{X} = 51.1726, \bar{Z} = 55.8611,$
   $\rho_{yx} = 0.3280, \quad \rho_{yx(2)} = 0.4770, \quad \rho_{yz} = 0.8460, \quad \rho_{yz(2)} = 0.7290,$
   $\rho_{xz} = 0.2970, \quad \rho_{xz(2)} = 0.5700, C_y = 0.1562, \quad C_{y(2)} = 0.1207,$
   $C_x = 0.0301, \quad C_{x(2)} = 0.0247, \quad C_z = 0.0586, \quad C_{z(2)} = 0.0541.$

2. Data set 2 [Source: Khare and Sinha (2012)]

   y: Number of literate persons in the village.
   x: Number of workers in the village.
   z: Number of non-workers in the village.

   $N = 109, n = 30, W_2 = 0.25, \bar{Y} = 145.30, \bar{X} = 165.26, \bar{Z} = 259.08,$
   $\rho_{yx} = 0.81, \quad \rho_{yx(2)} = 0.78, \quad \rho_{yz} = 0.90, \quad \rho_{yz(2)} = 0.87, \rho_{xz} = 0.81,$
   $\rho_{xz(2)} = 0.74, C_y = 0.76, C_{y(2)} = 0.68, C_x = 0.68, \quad C_{x(2)} = 0.57,$
   $C_z = 0.76, \quad C_{z(2)} = 0.54.$

3. Data set 3 [Source: Khare and Sinha (2009)]

   y: Number of agricultural labors in the village.
   x: Area (in hectares) of the village.
   z: Number of cultivators in the village.

$N = 96, n = 30, W_2 = 0.25, \bar{Y} = 137.92, \bar{X} = 144.87, \bar{Z} = 185.21,$
$\rho_{yx} = 0.77, \quad \rho_{yx(2)} = 0.72, \quad \rho_{yz} = 0.78, \rho_{yz(2)} = 0.78, \rho_{xz} = 0.81,$
$\rho_{xz(2)} = 0.72, C_y = 1.32, \quad C_{y(2)} = 2.08, C_x = 0.81, \quad C_{x(2)} = 0.94,$
$C_z = 1.05, \quad C_{z(2)} = 1.48.$

We use the following expression to obtain the percent relative efficiency (*PRE*) for different estimators using different values of h:

$$PRE = \frac{V(\bar{y}^*)}{MSE(i) \quad \text{or} \quad MSE(i)_{min}} \times 100, \qquad i = \bar{y}^*, \bar{y}^*_{Rao(P)}, \bar{y}^*_{Rao(P)}, \bar{y}^*_R, \ldots, \bar{y}^*_{prop}.$$

Results based on three data sets are given in Tables 2, 3 and 4.

Table 2: PRE of estimators with respect to $\bar{y}^*$ for data set 1

Estimator	h=2	h=4	h=6	h=8	h=16
$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000
$\bar{y}^*_R$	111.204	112.946	113.987	114.678	116.057
$\bar{y}^*_P$	84.9820	83.8472	83.200	82.782	81.974
$\bar{y}^*_{Rao(R)}$	107.908	105.704	104.460	103.662	102.134
$\bar{y}^*_{Rao(P)}$	88.1636	91.004	92.745	93.922	96.314
$\bar{y}^*_{exp(R)}$	106.369	107.189	107.672	107.991	108.621
$\bar{y}^*_{exp(P)}$	92.6901	92.032	91.654	91.407	90.929
$\bar{y}^*_{SK(R1)}$	112.749	114.116	114.927	115.465	116.532
$\bar{y}^*_{SK(P)}$	72.8894	74.829	76.007	76.799	78.397
$\bar{y}^*_{KB}$	114.506	117.822	119.946	121.417	124.495
$\bar{y}^*_{SK(R2)}$	114.819	118.348	120.506	121.961	124.915
$\bar{y}^*_{SK(d)}$	114.819	118.348	120.506	121.961	124.915
$\bar{y}^*_{prop}$	309.222	270.592	254.538	245.755	231.530

Table 3: PRE of estimators with respect to $\bar{y}^*$ for data set 2

Estimator	h=2	h=4	h=6	h=8	h=16
$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000
$\bar{y}^*_R$	277.329	269.560	265.571	263.144	258.764
$\bar{y}^*_P$	31.268	31.832	32.144	32.341	32.712
$\bar{y}^*_{Rao(R)}$	203.460	155.016	137.471	128.411	114.442
$\bar{y}^*_{Rao(P)}$	36.190	44.8313	51.411	56.588	69.561
$\bar{y}^*_{exp(R)}$	205.995	201.434	199.072	197.627	195.006
$\bar{y}^*_{exp(P)}$	52.514	53.144	53.488	53.705	54.111
$\bar{y}^*_{SK(R1)}$	90.356	112.140	128.782	141.911	174.939
$\bar{y}^*_{SK(P)}$	16.087	19.056	21.147	22.701	26.275
$\bar{y}^*_{KB}$	282.247	273.501	269.048	266.350	261.505
$\bar{y}^*_{SK(R2)}$	282.309	273.589	269.133	266.427	261.559
$\bar{y}^*_{SK(d)}$	282.309	273.589	269.133	266.427	261.559
$\bar{y}^*_{prop}$	549.804	521.751	510.740	505.259	498.002

Table 4: PRE of estimators with respect to $\bar{y}^*$ for data set 3

Estimator	h=2	h=4	h=6	h=8	h=16
$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000
$\bar{y}^*_R$	204.333	192.089	188.197	186.285	183.458
$\bar{y}^*_P$	47.615	50.484	51.556	52.117	52.991
$\bar{y}^*_{Rao(R)}$	142.598	118.102	111.493	108.419	104.068
$\bar{y}^*_{Rao(P)}$	59.014	73.728	80.668	84.707	91.670
$\bar{y}^*_{exp(R)}$	149.031	143.344	141.481	140.556	139.175
$\bar{y}^*_{exp(P)}$	67.732	70.041	70.875	71.305	71.967
$\bar{y}^*_{SK(R1)}$	170.523	175.322	177.041	177.925	179.283
$\bar{y}^*_{SK(P)}$	31.343	39.367	43.181	45.410	49.267
$\bar{y}^*_{KB}$	222.273	213.479	211.130	210.092	208.730
$\bar{y}^*_{SK(R2)}$	226.014	216.679	213.634	212.124	209.873
$\bar{y}^*_{SK(d)}$	226.014	216.679	213.634	212.124	209.873
$\bar{y}^*_{prop}$	289.857	289.956	290.856	291.516	292.816

5.1. Discussion and Findings

In Table 2, PRE of the proposed class of estimators $\bar{y}^*_{prop}$ and Rao (1991) ratio estimator $\bar{y}^*_{Rao(R)}$ decrease as the values of h increases from 2 to 16. On the other hand, the situation is reverse for the estimates $\bar{y}^*_R, \bar{y}^*_{exp(R)}, \bar{y}^*_{SK(R1)}, \bar{y}^*_{KB}, \bar{y}^*_{SK(R2)}, \bar{y}^*_{SK(d)}$.
In Table 3, the performances of the proposed estimator $\bar{y}^*_{prop}$ and all the other considered estimators decrease with an increase in the value of h.
In Table 4, PRE of the proposed class of estimators $\bar{y}^*_{prop}$ and Singh and Kumar (2008) estimator $\bar{y}^*_{SK(R1)}$, increase with an increase in the values of h. Also in this table, PRE of other estimators $\bar{y}^*_{Rao(R)}, \bar{y}^*_{exp(R)}, \bar{y}^*_{SK(R1)}, \bar{y}^*_{KB}, \bar{y}^*_{SK(R2)}, \bar{y}^*_{SK(d)}$ decreases with an increase in the value of h.
In Tables 2, 3 and 4, we observe that the product type estimators $\bar{y}^*_{Rao(P)}$ $\bar{y}^*_P, \bar{y}^*_{exp(P)}$ and $\bar{y}^*_{SK(P)}$ perform very poorly because of positive correlation in data sets 1, 2 and 3. Generally, we can use product type estimators when there exists a negative correlation between the study variable and the auxiliary variable.

6. Conclusion

We proposed a generalized class of estimators for estimating the population mean using information on two auxiliary variables under non-response in simple random sampling. Expressions for bias and MSE of the proposed generalized class of estimators are derived up to the first degree of approximation. The proposed estimator $\bar{y}^*_{prop}$ is compared with Hansen and Hurwitz (1946) estimator and other considered estimators. A numerical study is carried out to support the theoretical results. In Tables 2, 3 and 4, the proposed class of estimators performs better than all other competitor estimators under non-response in simple random sampling. The product type estimators perform poorly because of positive correlation in all data sets. Therefore, the proposed class of estimators $\bar{y}^*_{prop}$ is preferable in different situations, i.e. when no auxiliary variable, single auxiliary variable, and two auxiliary variables are used. It is observed that the Singh and Kumar (2008) estimators $\bar{y}^*_{SK(R1)}$ and $\bar{y}^*_{SK(d)}$ perform equally but $\bar{y}^*_{SK(d)}$ is preferable because of unbiasednes. All product type estimators perform poorly due to weak correlation between the study and the auxiliary variables.

Acknowledgment

Authors are thankful to referees for their valuable suggestions and comments, which led to the improvement of this article.

REFERENCES

BAHL, S., TUTEJA, R, K., (1991). Ratio and product type exponential estimator. Information and Optimization Sciences. 12, pp. 159–163.

EL-BADRY, M. A., (1956). A sampling procedure for mailed questionnaires. J. Amer. Statist. Assoc. 51, pp. 209–227.

HANSEN, M. H., HURWITZ, W. N., (1946). The problem of non-response in sampling surveys. Journal of American Statistical Association. 41, pp. 517–529.

ISMAIL, M., SHAHBAZ, M. S., HANIF, M., (2011). A general class of estimator of population mean in presence of non-response. Pakistan Journal of Statistics. 27 (4), pp. 467–476.

KHARE, B. B., SRIVASTAVA, S. R., (1981). A generalized regression ratio estimator for the population mean using two auxiliary variables. The Aligarh Journal of Statistics. 1(1), pp. 43–51.

KHARE, B. B., SRIVASTAVA, S., (1993). Estimation of population mean using auxiliary character in presence of non-response. Nat. Acad. Sci. Lett. India. 16, pp. 111–114.

KHARE, B. B., SRIVASTAVA, S., (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response. Nat. Acad. Sci. Lett. India. 65, pp. 195–203.

KHARE, B. B., SRIVASTAVA, S., (1997). Transformed ratio type estimators for the population mean in presence of non-response. Communication in Statistics Theory and Method. 17, pp. 1779–1791.

KHARE, B. B., SRIVASTAVA, S., (2010). Generalized two phase estimators for the population mean in the presence of non-response. The Aligarh journal of Statistics. 30, pp. 39–54.

KHARE, B. B., SINHA, R, R., (2007). Estimation of the ratio of the two population means using multi auxiliary characteristics in the presence of non-response. In Statistical Techniques in Life Testing, Reliability, Sampling Theory and Quality Control. 1, pp. 63–171.

KHARE, B. B., SINHA, R, R., (2009). On the class of estimators for population mean using multi auxiliary characters in the presence of non-response. Statistics in Transition-New Series. 10 (1), pp. 3–14.

Khare, B. B., Sinha, R, R., (2012). Improved classes of estimators for ratio of two means with with double sampling the non-respondents. Statistica. 49 (3), pp. 75–83.

KUMAR, S., BHOUGAL, S., (2011). Estimation of the population mean in presence of non-response. Communications of the Korean Statistical Society. 18 (4), pp. 537–548.

MUNEER, S., SHABBIR, J, KHALIL, A., (2017). Estimation of finite population mean in simple random sampling and stratified random sampling using two auxiliary variables. Communication in Statistics Theory and Methods. 46 (5), pp. 2181–2192.

RAO, P. S. R. S., (1986). Ratio estimation with sub sampling the non-respondents. Survey Methodology. 12 (2), pp. 217–230.

OLKIN, I., (1958). Multivariate ratio estimation for finite populations. Biometrika, 45, pp. 154–165.

SINGH, H. P., KUMAR, S., (2008). A regression approach to the estimation of the finite population mean in the presence of non-response. Aust. N. Z. J. Stat. 50(4), pp. 395–408.

SINGH, H. P., KUMAR, S., KOZAK, M., (2010) Improved estimation of finite population mean using sub-sampling to deal with non response in two phase sampling scheme. Communication in Statistics Theory and Method. 39 (5), pp. 791–802.

SRINATH, K. P., (1971). Multiphase sampling in non-response problems. Journal of American Statistical Association. 66, pp. 583–586.

SHABBIR, J., NASIR, S., (2013). On estimating the finite population mean using two auxiliary variables in two phase sampling in the presence of non response. Communication in Statistics-Theory and Methods. 42, pp. 4127–4145.

SHUKLA, D., PATHAK, S., THAKUR, N. S., (2012) A transformed estimator for estimation of population mean with missing data in sample surveys. Journal of current engineering research. 2 (1):50-55.

SINGH, V. K., SHUKLA, D., (1993) An efficient one parameter family of factor-type estimator in sample survey. Metron. 51 (1-2), pp. 139–159.

TABASUM, R., KHAN, I. A., (2006). Double sampling for ratio estimator for the population mean in the presence of non-response. Assam Statistical Review 20, pp. 73–83.