ESTIMATION OF POPULATION MEAN USING TRANSFORMED AUXILIARY VARIABLE AND NON-RESPONSE

Vishwantra Sharma and Sunil Kumar[1]
Department of Statistics, University of Jammu, India

ABSTRACT
In survey sampling, the errors which are mostly studied during estimation are sampling errors. However, the properties of estimators are more influenced by non-sampling errors than sampling errors. This paper addresses the problem of estimating the finite population mean of the study variable y using auxiliary information in sample surveys in the presence of non-response. We have proposed an estimator for estimating the population mean of study variable when the parameter of auxiliary variable x is known. The bias and mean squared error (MSE) of the proposed estimator are obtained to the first degree of approximation. The minimum mean square error of the proposed estimator is also obtained. A Simulation study has been carried out to support the theoretical results. The comparison of the proposed estimator with other estimators is made to show that our proposed estimator is more efficient than the other estimators in terms of percent relative efficiency (PRE).

KEYWORDS: Study variable, Auxiliary Variable, Bias, Mean squared error, Non-response, Simulation.

MSC: 62D05

RESUMEN
En las encuestas por muestreo, los errores que han sido más estudiados en la estimación son los de muestreo. Sin embargo las propiedades delos estimadores son más influenciados por los errores ajenos al muestreo que por los de muestreo. Este paper trata del problema de estimar la media poblacional de la variable de estudioy cuando hay informaciónauxiliar en el muestro, en la presencia de no-respuestas. Nosotrosproponemos un estimador para estimar la media poblacional de la variable de estudio cuando el parámetro de la variable auxiliar x es conocida. Son obtenidos el sesgo y el error cuadrático medio (MSE) del propuesto estimadorhasta el primer grado de aproximación. El MSE mínimo del estimador propuesto también es obtenido. Un estudio de Simulacións desarrollado para soportar los resultados teoréticos. La comparación del estimador propuesto con otros es desarrollada para mostrar que este es más eficiente que los otros en términos del porciento de la eficiencia relativa (PRE).

PALABRAS CLAVE: Variable de Estudio, Variable Auxiliar ,Sesgo, Error Cuadrático Medio, No-Respuesta, Simulación.

1. INTRODUCTION

A surveying error arises whenever there is a discrepancy between the reality and statements. Usually, there are two types of error sampling errors and non-sampling errors. Sampling errors arise when the selected sample does not contain the true characteristics of the population. Non-sampling error arises due to several reasons like an error in the questionnaire, information provided by respondents, data collection and preparation.

Survey often yields estimates that are subjected to bias because of either non-response or measurement error. In sample survey non-response arises when one fails to get information from some units of the population due to various reasons such as when someone refuses to answer, sometimes the respondents are not available and there is possibility that the respondents didn't get the question due to lack of interest or unable to understand what has been asked in questionnaire. Thus, the researcher needs to be extra cautious while designing the questionnaire for the survey so that such errors should be minimized.

Various methods are available for eliminating non-response bias (e.g. re-weighting the data, individual-level mode of selection or deriving bounds on the true population parameter). Mostly, the information is not available or obtained from all the units during surveys so non-response problems may creep into the study from the very beginning.

In sample surveys, the problem of non-response is common in mail surveys than in personal interviews. To overcome this problem of non-response the researcher has to approach the non-respondent and obtained the

[1]Email: vishwantrasharma07@gmail.com, sunilbhougal06@gmail.com

information. Hansen and Hurwitz [4] firstly deal with the problem of non-response, they proposed a sampling scheme that involves taking a subsample of non-respondents after the mail survey and the obtain information from personal interview. These surveys have the benefit that the data can be collected in a relatively inexpensive way. Cochran [2], Rao [14], Khare and Srivastava [7, 8], Singh and Kumar [12], Kumar *et al*. [13], Kumar and Bhougal [11], Kumar [9] and Kumar [10] studied the problem of non-response in sample surveys using auxiliary information. Gupta and Shabbir [3] have suggested a general class of ratio estimators with known auxiliary information. Shabbir and Gupta [15] define ratio-type exponential estimator using a transformed auxiliary variable. Further, Kaur [6] proposed a generalized class of ratio-type exponential estimators of population mean under the general linear transformation of auxiliary variables.

In this paper, we proposed an exponential type of study variable by using a very general linear transformation of the available auxiliary variable in the presence of non-response on both study as well as an auxiliary variable. We estimate the population mean $\bar{Y}$ of study variable intending to minimize the bias and MSE of the proposed estimator using auxiliary information. The MSE is minimized and the simulation study is performed to take an overview about the PRE (percent relative efficiency) of proposed estimator with already existing estimators.

2. SAMPLING PROCEDURE AND SOME WELL-DEFINED ESTIMATORS

Let a sample of size n is selected from the population of size N by simple random sampling without replacement (SRSWOR). Let y and x be the study and the auxiliary variable respectively. Let $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$,$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ and $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$ be the population mean and variance of the study variable and auxiliary variable, respectively. The problem of non-response is common and more widespread in mail surveys than in a personal interview. Hansen and Hurwitz (1946) were the first one to deal with this situation and proposed methodology for the estimation of population mean of the study variable 'y'. They proposed a double sampling scheme for estimating population mean, where a simple random sample of size 'n' is selected and the questionnaire is mailed to the sampled units; n_1 be the number of respondents and n_2 be the non-respondents in the sample, and a subsample of size $r(=\frac{n_2}{k}; k > 1)$ is taken from the non-respondents, where k is the inverse sampling ratio. They considered the situation where the population of size N is composed of two mutually exclusive groups, the N_1 respondents and the $N_2(= N - N_1)$ non-respondents.

Let $\bar{Y}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} y_i$ and $S_{y_1}^2 = \frac{1}{N_1-1}\sum_{i=1}^{N_1}(y_i - \bar{Y}_1)^2$denotes the mean and the variance of the response group, respectively, and let $\bar{Y}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} y_i$ and $S_{y(2)}^2 = \frac{1}{N_2-1}\sum_{i=1}^{N_2}(y_i - \bar{Y}_2)^2$denote the mean and the variance of the non-response group, respectively of study variable.

The population mean can be written as$\bar{Y} = W_1\bar{Y}_1 + W_2\bar{Y}_2$, where $W_1 = \frac{N_1}{N}$ and $W_2 = \frac{N_2}{N}$.

The usual unbiased estimator for the population mean $\bar{\mathrm{Y}}$in the presence of non-response is given by

$$\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_2, \tag{1}$$

with variance

$$Var(\bar{y}^*) = \lambda S_y^2 + \theta S_{y_{(2)}}^2 \tag{2}$$

where $w_1 = \frac{n_1}{n}, w_2 = \frac{n_2}{n},\ \lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$ and $\theta = \frac{w_2(k-1)}{n}$.

Rao (1986) suggested ratio, product and regression estimator in the presence of non-response on the study as well as the auxiliary variable when population mean '$\bar{X}$' is known as

$$\bar{y}_R^* = \bar{y}^*\frac{\bar{X}}{\bar{x}^*} \tag{3}$$

$$\bar{y}_P^* = \bar{y}^*\frac{\bar{x}^*}{\bar{X}} \tag{4}$$

$$\bar{y}_{lr}^* = \bar{y}^* + b_{yx}(\bar{X} - \bar{x}^*) \tag{5}$$

The MSE of $\bar{y}_R^*$, $\bar{y}_P^*$ and $\bar{y}_{lr}^*$ to the first degree of approximation, as

$$MSE(\bar{y}_R^*) = \lambda\left(S_y^2 + R^2S_x^2 - 2RS_{yx}\right) + \theta\left(S_{y(2)}^2 + R^2S_{x(2)}^2 - 2RS_{yx(2)}\right) \tag{6}$$

$$MSE(\bar{y}_P^*) = \lambda\left(S_y^2 + R^2S_x^2 + 2RS_{yx}\right) + \theta\left(S_{y(2)}^2 + R^2S_{x(2)}^2 + 2RS_{yx(2)}\right). \tag{7}$$

$$MSE(\bar{y}_{lr}^*) = \lambda\left(1 - \rho_{yx}^2\right)S_y^2 + \theta\left(1 - \rho_{yx(2)}^2\right)S_{y(2)}^2 \tag{8}$$

where $\bar{\mathrm{x}}^* = w_1\bar{\mathrm{x}}_1 + w_2\bar{\mathrm{x}}_2, \bar{\mathrm{x}}_1 = \frac{1}{\mathrm{N}_1}\sum_{\mathrm{i}=1}^{\mathrm{N}_1}\mathrm{x}_\mathrm{i}$, $\bar{\mathrm{x}}_2 = \frac{1}{\mathrm{N}_2}\sum_{\mathrm{i}=1}^{\mathrm{N}_2}\mathrm{x}_\mathrm{i}, R\ = \frac{\bar{Y}}{\bar{X}}$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$,

$S_{y(2)}^2 = \frac{1}{N_2-1}\sum_{i=1}^{N_2}(y_i - \bar{Y})^2$, $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$, $S_{x(2)}^2 = \frac{1}{N_2-1}\sum_{i=1}^{N_2}(x_i - \bar{X})^2$,
$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})(x_i - \bar{X})$, $S_{yx(2)} = \frac{1}{N_2-1}\sum_{i=1}^{N_2}(y_i - \bar{Y}_2)(x_i - \bar{X}_2)$.
ρ_{yx} and $\rho_{yx(2)}$ are the correlation coefficient between the study and auxiliary variable for respondents and non-respondents, respectively.
Kadilar and Cingi [5] proposed a ratio regression type estimator for the population mean '$\bar{Y}$' in the presence of non-response on study and auxiliary variable as

$$\bar{y}_{KC}^* = \bar{y}^* + b_{yx}(\bar{X} - \bar{x}^*)\left(\frac{\bar{X}}{\bar{x}^*}\right) \tag{9}$$

and $MSE(\bar{y}_{KC}^*)$ is given by

$$MSE(\bar{y}_{KC}^*) = \lambda\{S_y^2 + S_x^2(1 - \rho_{yx}^2)\} + \theta\{S_{y(2)}^2 + S_{x(2)}^2(1 - \rho_{yx(2)}^2)\} \tag{10}$$

In the presence of non-response over study and auxiliary variables, Bahl and Tuteja [1] estimator becomes

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

and $MSE(\bar{y}_{BT}^*)$ is given by

$$MSE(\bar{y}_{BT}^*) = \lambda\left(S_y^2 + \frac{S_x^2}{4} + 2\rho_{yx}S_yS_x\right) + \theta\left(S_{y(2)}^2 + \frac{S_{x(2)}^2}{4} + 2\rho_{yx(2)}S_{y(2)}S_{x(2)}\right) \tag{12}$$

Further, in the next section, we have suggested a difference-cum-exponential estimator in the presence of non-response on study and auxiliary variables.

3. THE PROPOSED ESTIMATOR

Let ω(≠ 0) and ψ be either known or function of any known parameters of the auxiliary variable x such as standard deviation, coefficient of variation, coefficient of kurtosis, coefficient of skewness and coefficient of correlation, etc. In our estimator, we define the transformation of variable

$$Z = \omega X + \psi \tag{13}$$

So we have

$$\bar{Z} = \omega\bar{X} + \psi \text{ and } \bar{z} = \omega\bar{x} + \psi \tag{14}$$

where $\bar{Z}$ and $\bar{z}$ are the population and sample mean of the transformed auxiliary variable z, respectively. The value of $\bar{Z}$ is known to us as ω, ψ and $\bar{X}$ are assumed to be known.
By getting motivation from Gupta and Shabbir [3], Shabbir and Gupta [15] and Kaur [6] we propose the difference cum exponential estimator in the presence of non-response on study and auxiliary variable by using transformed auxiliary variable when population mean $\bar{X}$ is known, as

$$t = \{W_{1g}\bar{y}^* + W_{2g}(\bar{X} - \bar{x}^*)\}exp\left[\frac{\bar{z}^\delta - \bar{z}}{\bar{z}^\delta + \bar{z}}\right] \tag{15}$$

where $\bar{z}^\delta = \frac{N\bar{Z} - n\bar{z}^*}{N-n}$ and $\bar{z}^* = \omega\bar{x}^* + \psi$.
Let $\bar{y}^* = \bar{Y}(1+\epsilon)$; $\bar{x}^* = \bar{X}(1+\eta)$, also, $E(\epsilon) = E(\eta) = 0$, $E(\epsilon^2) = (\lambda S_y^2 + \theta S_{y(2)}^2) = \alpha_1$; $E(\eta^2) = (\lambda S_x^2 + \theta S_{x(2)}^2) = \alpha_2$; and $E(\eta\epsilon) = (\lambda\rho_{yx}S_yS_x + \theta\rho_{yx(2)}S_{y(2)}S_{x(2)}) = \alpha_3$
where $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right) = \frac{1-f}{n}$; and $f = \frac{n}{N}$; $\theta = \frac{W_2(k-1)}{2}$ and $W_2 = \frac{N_2}{N}$.
Expanding right hand side of equation (15) to the first degree of approximation, we get

$$t = W_{1g}\left\{\bar{Y} + \epsilon + \bar{Y}(A^{-2}\phi_1\eta^2) - \bar{Y}A^{-1}N\omega\eta - A^{-1}N\omega\eta\epsilon + \frac{\bar{Y}}{2}A^{-2}N^2\omega^2\eta^2\right\} - W_{2g}\{\eta - A^{-1}N\omega\eta^2\}; \tag{16}$$

Subtracting $\bar{Y}$on both sides of equation (16), we get

$$(t - \bar{Y}) = W_{1g}\left\{\bar{Y} + \epsilon + \bar{Y}(A^{-2}\phi_1\eta^2) - \bar{Y}A^{-1}N\omega\eta - A^{-1}N\omega\eta\epsilon + \frac{\bar{Y}}{2}A^{-2}N^2\omega^2\eta^2\right\} - W_{2g}\left\{\eta - A^{-1}N\omega\eta^2 + YW2g \right. \tag{17}$$

where $A = 2(N-n)(\psi + \omega\bar{X})$, $\phi_1 = N(N-2n)\omega^2$.
Taking expectation on both sides of equation (17), we will get the bias of the proposed estimator 't' as

$$B(t) = W_{1g}\left\{\bar{Y} + \bar{Y}A^{-2}\phi_1\alpha_2 - A^{-1}N\omega\alpha_3 + \frac{\bar{Y}}{2}A^{-2}N^2\omega^2\alpha_2\right\} + W_{2g}A^{-1}N\omega\alpha_2 - \bar{Y} \tag{18}$$

Squaring equation (17) up to first order of approximation, we get

$$(t-\bar{Y})^2 = \left[W_{1g}\left\{\bar{Y}+\epsilon+\bar{Y}(A^{-2}\phi_1\eta^2)-\bar{Y}A^{-1}N\omega\eta - A^{-1}N\omega\eta\epsilon+\frac{\bar{Y}}{2}A^{-2}N^2\omega^2\eta^2\right\} - W_{2g}\left\{-A^{-1}N\omega\eta^2+\frac{\bar{Y}}{W_{2g}}\right\}\right]^2$$

and taking expectation on both sides, one can obtain the MSE of 't' as

$$MSE(t) = W_{1g}^2\{\bar{Y}^2+\alpha_1+\phi_2^2\bar{Y}\alpha_2+2\bar{Y}\alpha_3+2\phi_2\phi_3\alpha_2\}+W_{2g}^2\alpha_2-2W_{2g}\phi_2\bar{Y}\alpha_2+\bar{Y}^2-2W_{1g}W_{2g}\{\alpha_3- 2\phi 2Y\alpha 2-2W1gYY-\phi 2\alpha 3+Y\phi 3\alpha 2 \tag{19}$$

where $\phi_2 = A^{-1}N\omega$ and $\phi_3 = A^{-1}\phi_1+\frac{\phi_2^2}{2}$.

To minimize MSE(t), we differentiate MSE(t) with respect to W_{1g} and $W_{2g}$and equate to zero, we get the optimum value of W_{1g} and W_{2g} say $\widehat{W}_{1g(opt)}$ and $\widehat{W}_{2g(opt)}$.

$$W_{1g} = \frac{\bar{Y}\alpha_2\{(1-\bar{Y})\phi_2\alpha_3-\bar{Y}\alpha_2(2\phi_2^2-\phi_3)\}}{\{B-\alpha_3-2\phi_2\bar{Y}\alpha_2\}} = \widehat{W}_{1g(opt)}(\text{say}) \tag{20}$$

$$W_{2g} = \frac{[\phi_2\bar{Y}\alpha_2\{B-\alpha_3-2\phi_2\bar{Y}\alpha_2\}+(\alpha_3-2\bar{Y}\phi_2\alpha_2)\bar{Y}\alpha_2\{(1-\bar{Y})\phi_2\alpha_3-\bar{Y}\alpha_2(2\phi_2^2-\phi_3)\}]}{[\alpha_2\{B-\alpha_3-2\phi_2\bar{Y}\alpha_2\}]} = \widehat{W}_{2g(opt)}(\text{say}) \tag{21}$$

where, $B = \alpha_2\{\bar{Y}+\alpha_1+\phi_2^2\bar{Y}\alpha_2+2\bar{Y}\alpha_3+2\phi_2\phi_3\alpha_2\}$.

On substituting the optimum value of $\widehat{W}_{1g(opt)}$ and $\widehat{W}_{2g(opt)}$ from equation (20) and (21) in (19), we get the minimum mean square error of the proposed estimator, as

$$minMSE(t) = \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)+\widehat{W}_{1g}^2\alpha_1+\alpha_2 L+\alpha_3 M, \tag{22}$$

where, $L = \widehat{W}_{1g}^2\phi_2^2\bar{Y}+\widehat{W}_{2g}^2-2\widehat{W}_{2g}\phi_2\bar{Y}+4\widehat{W}_{1g}\widehat{W}_{2g}\phi_2\bar{Y}-2\widehat{W}_{1g}\bar{Y}^2\phi_3$, $M = 2\big(\widehat{W}_{1g}^2\bar{Y}-\widehat{W}_{1g}^2\phi_2\phi_3- W1gW2g+W1g\phi 2Y$.

4. EFFICIENCY COMPARISON

For the efficiency comparison, we have compared $minMSE(t)$ of our proposed estimator with the MSEconsidered estimators which are defined in section 2 i.e. usual unbiased estimator $\bar{y}^*$, Rao's [14] ratio estimator ($\bar{y}_R^*$), product estimator ($\bar{y}_P^*$) and regression estimator ($\bar{y}_{lr}^*$), estimator proposed by Kadilar and Cingi [5] $\bar{y}_{KC}^*$ and Bahl and Tuteja [1] $\bar{y}_{BT}^*$. From (2), (5), (6), (8), (10), (12) and (22), we have

a. $\mathrm{Var}(\bar{y}^*) - minMSE(t) \geq 0$
if $\alpha_1\left(1-\widehat{W}_{1g}^2\right)-\alpha_2 L-\alpha_3 M \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$
b. $MSE(\bar{y}_R^*) - minMSE(t) \geq 0$
if $\alpha_1\left(1-\widehat{W}_{1g}^2\right)+\alpha_2(R^2-L)+\alpha_3(2R-M) \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$
c. $MSE(\bar{y}_P^*) - minMSE(t) \geq 0$
if $\alpha_1\left(1-\widehat{W}_{1g}^2\right)+\alpha_2(R^2-L)+\alpha_3(2R-M) \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$
d. $MSE(\bar{y}_{lr}^*) - minMSE(t) \geq 0$
if $\alpha_4-\widehat{W}_{1g}^2\alpha_1-\alpha_2 L-\alpha_3\mathrm{M} \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$
where $\alpha_4 = \lambda\left(1-\rho_{yx}^2\right)+\theta\left(1-\rho_{yx(2)}^2\right)$
e. $MSE(\bar{y}_{KC}^*) - minMSE(t) \geq 0$
if $\alpha_1\left(1-\widehat{W}_{1g}^2\right)+\alpha_2(1-L)-\alpha_3 M-\alpha_5 \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$
where $\alpha_5 = \lambda\rho_{yx}^2 S_x^2+\theta\rho_{yx(2)}^2 S_{x(2)}^2$
f. $MSE(\bar{y}_{BT}^*) - minMSE(t) \geq 0$
if $\alpha_1\left(1-\widehat{W}_{1g}^2\right)+\alpha_2\left(\frac{1}{4}-L\right)+\alpha_3(2-M) \geq \bar{Y}^2\left(1+\widehat{W}_{1g}^2\right)$ (23-28)

If all the conditions defined in equation (23-28) are true, then we can say that our proposed estimator is efficient than considered estimators. Next section, we show the practical performance of our proposed theoretical results.

5. THE SIMULATION STUDY

In simulation, we have generated a population of size N from normal distribution in whichN_1 are respondents and $N_2$are non-respondents.We take sample of size n in which sample of size $n_1$is taken from respondents and sample of size $n_2 = (n-n_1)$ is taken from non-respondents. $X = rnorm(N,0,1)$ is the auxiliary variable and $Y = X + rnorm(N,0,1)$ is the study variable. We have generated 4 populations and values of MSEs for different value of k are provided in tables$(1-4)$. For comparison we have taken two cases, in first case ψ and ω takes the constant value and in second case ψ and ω takes parametric values for different values of k.

Population I: $X = rnorm(N, 0, 1); Y = X + rnorm(N, 0,1); N = 2000;\ N_1 = 1500;$
$N_2 = 500;$ n $= 800;\ n_1 = 500;\ n_2 = 300.$

Table 1: Mean Square Error (MSE) and PRE of the estimators for population I.

For $\psi = 1$ and $\omega = 1$							
1/k	$\bar{y}^*$	$\bar{y}_R^*$	$\bar{y}_P^*$	$\bar{y}_{lr}^*$	$\bar{y}_{KC}^*$	$\bar{y}_{BT}^*$	t
1/2	0.003038	0.003696	0.003649	0.003038	0.004638	0.003496	0.0003537
	100.00	82.21	83.26	100.00	65.78	88.30	859.11
1/3	0.003646	0.004478	0.004389	0.003649	0.005558	0.004141	0.0003537
	100.00	81.43	83.08	99.93	65.60	88.07	1030.98
1/4	0.004255	0.005261	0.005128	0.004259	0.006498	0.004839	0.0003537
	100.00	80.88	82.96	99.89	65.47	87.91	1202.84
1/5	0.004863	0.0180451	0.0172728	0.0051576	0.0075647	0.0057203	0.0003537
	100.00	80.47	82.86	99.86	65.37	87.79	1374.68
For $\psi = \beta_2(x)$and $\omega = cov(X)$							
1/2	0.003038	0.003696	0.003649	0.003038	0.004638	0.003496	0.0003534
	100.00	82.21	83.26	100.00	65.78	88.30	859.69
1/3	0.003646	0.004478	0.004389	0.003649	0.005558	0.004141	0.0003534
	100.00	81.43	83.08	99.93	65.60	88.07	1031.74
1/4	0.004255	0.005261	0.005128	0.004259	0.006498	0.004839	0.0003534
	100.00	80.88	82.96	99.89	65.47	87.91	1203.78
1/5	0.004863	0.0180451	0.0172728	0.0051576	0.0075647	0.0057203	0.0003534
	100.00	80.47	82.86	99.86	65.37	87.79	1375.82

Population II : $X = rnorm(N, 0, 1); Y = X + rnorm(N, 0,1); N = 4000;\ N_1 = 3000;$
$N_2 = 1000;$ n $= 2000;\ n_1 = 1500;\ n_2 = 500.$

Table 2: Mean Square Error (MSE) and PRE of the estimators for population II.

For $\psi = 1$ and $\omega = 1$							
1/k	$\bar{y}^*$	$\bar{y}_R^*$	$\bar{y}_P^*$	$\bar{y}_{lr}^*$	$\bar{y}_{KC}^*$	$\bar{y}_{BT}^*$	t
1/2	0.001229	0.004825	0.004734	0.001245	0.004638	0.001401	0.0000667
	100.00	25.48	25.97	98.71	65.78	87.81	1842.28
1/3	0.001475	0.005899	0.005714	0.003649	0.001508	0.001694	0.0000667
	100.00	25.01	25.82	99.93	97.84	87.11	2210.81
1/4	0.001721	0.006971	0.006694	0.004259	0.001770	0.001987	0.0000667
	100.00	24.69	25.71	99.89	97.24	86.62	2579.35
1/5	0.001967	0.008040	0.007675	0.0051576	0.002033	0.002281	0.0000667
	100.00	24.46	25.62	99.86	96.75	86.23	2947.87
For $\psi = \beta_2(x)$ and $\omega = cov(X)$							
1/2	0.001229	0.004825	0.004734	0.001245	0.004638	0.001401	0.0000667
	100.00	25.48	25.97	98.71	65.78	87.81	1842.28
1/3	0.001475	0.005899	0.005714	0.003649	0.001508	0.001694	0.0000667
	100.00	25.01	25.82	99.93	97.84	87.11	2210.81
1/4	0.001721	0.006971	0.006694	0.004259	0.001770	0.001987	0.0000667
	100.00	24.69	25.71	99.89	97.24	86.62	2579.35
1/5	0.001967	0.008040	0.007675	0.0051576	0.002033	0.002281	0.0000667
	100.00	24.46	25.62	99.86	96.75	86.23	2947.87

Population III : $X = rnorm(N, 0, 1); Y = X + rnorm(N, 0,1); N = 8000;\ N_1 = 5500;$
$N_2 = 2500;$ n $= 3000;\ n_1 = 2000;\ n_2 = 1000.$

Table 3: Mean Square Error (MSE) and PRE of the estimators for population III.

For $\psi = 1$ and $\omega = 1$							
1/k	$\bar{y}^*$	$\bar{y}_R^*$	$\bar{y}_P^*$	$\bar{y}_{lr}^*$	$\bar{y}_{KC}^*$	$\bar{y}_{BT}^*$	t
1/2	0.0008815	0.000908	0.000907	0.000873	0.001320	0.000985	0.00007745
	100.00	97.03	97.17	100.86	66.77	89.40	1138.05
1/3	0.001091	0.001120	0.001118	0.001076	0.001627	0.001215	0.00007745

	100.00	97.39	97.60	101.38	67.05	89.81	1409.03
1/4	0.001301	0.001332	0.001329	0.001279	0.001935	0.001444	0.00007745
	100.00	97.63	97.90	101.74	67.23	90.09	1680.01
1/5	0.001511	0.001545	0.001540	0.001481	0.002243	0.001673	0.00007745
	100.00	97.81	98.11	102.37	67.37	90.29	1950.98
For $\psi = \beta_2(x)$ and $\omega = cov(X)$							
1/2	0.0008815	0.000908	0.000907	0.000873	0.001320	0.000985	0.00007744
	100.00	97.03	97.17	100.86	66.77	89.40	1138.05
1/3	0.001091	0.001120	0.001118	0.001076	0.001627	0.001215	0.00007744
	100.00	97.39	97.60	101.38	67.05	89.81	1409.29
1/4	0.001301	0.001332	0.001329	0.001279	0.001935	0.001444	0.00007744
	100.00	97.63	97.90	101.74	67.23	90.09	1680.34
1/5	0.001511	0.001545	0.001540	0.001481	0.002243	0.001673	0.00007744
	100.00	97.81	98.11	102.37	67.37	90.29	1951.39

Population IV : $X = rnorm(N, 0, 1); Y = X + rnorm(N, 0,1); N = 25000;\ N_1 = 18000;$ $N_2 = 7000;$ n $= 10000;\ n_1 = 6000;\ n_2 = 4000.$

Table 4: Mean Square Error (MSE) and PRE of the estimators for population IV.

For $\psi = 1$ and $\omega = 1$							
1/k	$\bar{y}^*$	$\bar{y}_R^*$	$\bar{y}_P^*$	$\bar{y}_{lr}^*$	$\bar{y}_{KC}^*$	$\bar{y}_{BT}^*$	t
1/2	0.0002566	0.0003342	0.0003318	0.0002571	0.0003854	0.0002895	0.000020024
	100.00	76.78	77.34	99.80	66.58	88.62	1281.69
1/3	0.0003127	0.0004095	0.0004046	0.0003138	0.00047	0.000353	0.000020024
	100.00	76.38	77.29	99.67	66.67	88.46	1562.07
1/4	0.0003689	0.0004847	0.0004775	0.0003705	0.0005545	0.0004175	0.000020024
	100.00	76.10	77.25	99.57	66.52	88.35	1842.454
1/5	0.000425	0.000560	0.0005504	0.0004271	0.0006391	0.0004815	0.000020024
	100.00	75.90	77.23	99.51	66.50	88.27	2122.83
For $\psi = \beta_2(x)$ and $\omega = cov(X)$							
1/2	0.0002566	0.0003342	0.0003318	0.0002571	0.0003854	0.0002895	0.000020023
	100.00	76.78	77.34	99.80	66.58	88.62	1281.74
1/3	0.0003127	0.0004095	0.0004046	0.0003138	0.00047	0.000353	0.000020023
	100.00	76.38	77.29	99.67	66.67	88.46	1562.13
1/4	0.0003689	0.0004847	0.0004775	0.0003705	0.0005545	0.0004175	0.000020023
	100.00	76.10	77.25	99.57	66.52	88.35	1842.53
1/5	0.0004250	0.000560	0.0005504	0.0004271	0.0006391	0.0004815	0.000020023
	100.00	75.90	77.23	99.51	66.50	88.27	2122.92

From the tables$(1 - 4)$it is noted that our proposed estimator is performing better than the usual unbiased estimator and other considered estimators for each population in which ψ and ω takes constant and parametric values, respectively. We can observed from table 1, 3 and 4 that with the increase in the value of k, the variance and MSE of the considered estimators increases where as PRE of the considered estimator decreases but for the proposed estimator MSE remains same and PRE increases. From table 2, it can be seen that MSE and PRE follow same pattern as in other tables for the considered estimators but for proposed estimators the value of MSE and PRE are same. Also, the MSE of our proposed estimator$t$is very small among the other considered estimators viz. $\bar{y}_R^*,\ \bar{y}_p^*,\ \bar{y}_{lr}^*,\ \bar{y}_{KC}^*$ and $\bar{y}_{BT}^*$. From the table, it can be seen that the results for both the cases are almost same and there is minute difference in the MSE of our proposed estimator for case 1 and case 2 in population I and II.

6. CONCLUSION

This article considers the problem of estimating the population mean of the study variable by using auxiliary information in the presence of non-response. We have suggested a difference-cum-exponential type estimator by introducing a transformation on auxiliary variable and studied its properties. We compare our proposed estimator with usual unbiased estimator $\bar{y}^*$, ratio estimator $\bar{y}_R^*$, product estimator $\bar{y}_P^*$, regression estimator $\bar{y}_{lr}^*$, ratio-regression estimator $\bar{y}_{KC}^*$ and $\bar{y}_{BT}^*$ and developed the conditions under which our proposed estimator 't' is efficient. Next, we studied the simulation and obtained the mean square error (MSE) and percent relative efficiency (PRE) of different estimators by considering four different populations to see the performance of the estimators in different situations. With simulation results, we find that the results support the theoretical results obtained in section 5 i.e. the proposed estimator is having the minimum MSE as compared to other considered estimators. Also, from the PRE results we can say that proposed estimator has larger efficiency than the considered estimators. Thus, we recommend our proposed estimator in practice under the situations considered in the simulation study.

Acknowledgement: We thank the learned reviewers for their thorough review and highly appreciate the comments and suggestions, which significantly improved the quality of the paper.

RECEIVED: JUNE, 2019
REVISED: OCTOBER, 2019

REFERENCES

[1] BAHL, S. and TUTEJA, R. K. (1991): Ratio and product type exponential estimator. **Intro. Optimiz. Sci.,** 12, 159-163.

[2] COCHRAN, W.G.(1977): **Sampling Techniques**. Third Edition. New York: John Wiley & Sons, Inc.

[3] GUPTA, S., and SHABBIR, J.(2008): On estimating in estimating the population mean in simple random sampling. **J. Appl. Stat.** 35, 559–566.

[4] HANSEN, M.H. and HURWITZ, W.N. (1946): The problem of non-response in sample surveys. **J. Amer. Statist. Assoc.**, 41, 517-529.

[5] KADILAR, C. and CINGI, H. (2004): Ratio estimators in simple random sampling. **Applied Mathematics and Computation**, 151, 893-902.

[6] KAUR, P. (2013): Some contributions to efficient estimation of population parameters using auxiliary information. **Chapter 6, Ph.D Thesis, Department of Mathematics, Guru Nanak Dev University, Amritsar, India**.

[7] KHARE, B.B. and SRIVASTAVA, S. (1993): Estimation of population mean using auxiliary character in presence of non-response. **National Academy Science Letters.**16, 111-114.

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

[9] KUMAR, S. (2012): Estimation of population ratio in presence of non-response in successive sampling. **Journal of Statistical Theory and Applications,** 11, 293-310.

[10] KUMAR, S. (2016): Improved estimation of population mean in presence of non-response and measurement error. **Journal of Statistical Theory and Practice**, 10, 707-720.

[11] KUMAR, S. and BHOUGAL, S. (2011): Estimation of the population mean in presence of non-response. **Commun. Korean Statist. Soc.,** 18, 1-12.

[12] SINGH, H.P. and KUMAR, S. (2008): A regression approach to the estimation of the finite population mean in the presence of non-response. **Australian & New Zealand Journal of Statistics**. 50, 395 - 408.

[13] KUMAR, S., SINGH, H.P., BHOUGAL, S. and GUPTA, R. (2011): A class of ratio-cum-product type estimators under double sampling in the presence of non-response. **Hacettepe J. Math. and Statist.,** 40, 589-599.

[14] RAO, P.S.R.S. (1986): Ratio estimation with sub sampling the non-respondents. **Survey Methodology**, 12, 217-230.

[15] SHABBIR, J. and GUPTA, S. (2011): On estimating finite population mean in simple and stratified random sampling. **Communications in Statistics-Theory and Methods**, 40(2), 199-212.