Volume 70, Issue 3, 2026

Journal of Scientific Research

of

The Banaras Hindu University

Estimating The Finite Population Mean in Systematic Sampling Under Non-Response

Manoj K. Chaudhary[*1], Anil Prajapati[2] and Basant K. Ray[3]

***Abstract:* The present paper aims to estimate the population mean in systematic sampling utilizing the information on a single auxiliary variable under non-response/missing observations. We have considered the situation of non-response/missing observations in which certain observations on some sampling units for either study variable or auxiliary variable are missing. In this connection, we have suggested some ratio, product and regression-type estimators of the population mean. A general class of estimators of the population mean has also been suggested. The fundamental properties of the suggested ratio, product and regression-type estimators have been discussed in detail. The expressions for the bias and MSE of the suggested class of estimators have been derived up to the first order of approximation. The optimum estimator of the suggested class has also been pioneered out. The theoretical results have been demonstrated through some empirical data.**

***Index Terms:* Population mean, class of estimators, systematic sampling, auxiliary information, non-response/missing observations.**

I. Introduction

The systematic sampling scheme is widely used and most applicable technique to select a sample from the population. In this sampling scheme, the first unit is randomly selected and rest of the units is being automatically selected according to some predetermined rule involving the regular spacing of the units. Madow and Madow (1944), Madow (1946), Cochran (1946) and Lahiri (1954) have discussed the systematic sampling in details. Yates (1948) and Buckland (1951) have given the reviews of the work made in the systematic sampling. Hasel (1942), Finney (1948) and Nair and Bhargava (1951) have discussed the applications of systematic sampling in forest survey. Sukhatme *et al.* (1958) have demonstrated the application of systematic sampling in estimating the catch of fish. Kushwaha and Singh (1989), Singh and Singh (1998), Singh *et al.* (2012), Khan and Singh (2015) and Mishra *et al.* (2018) have utilized the auxiliary information in developing the estimators of the population parameters under systematic sampling.

The above works are discussed under the situation where the population is free from the non-response or incompleteness. But, in a mail survey, it is not advisable to avoid the occurrence of non-response and it might play an important role in determining the precision of the estimates. In such a situation, Hansen and Hurwitz (1946) introduced a technique of sub-sampling of non-respondents to handle the problem of non-response. This technique is commonly used in estimating the parameters of the population. Toutenburg and Srivastava (1998) have suggested a new method of estimating the ratio of population means in survey sampling under some missing observations. Grover and Kaur (2014) have proposed some exponential ratio-type estimators of the population mean under non-response/missing observations. Chaudhary *et al.* (2012), Chaudhary *et al.* (2013) and Shahzad *et al.* (2017) have proposed some families of estimators of the population mean in systematic sampling under non-response.

In the present paper, we have suggested some new methods of estimating the population mean in systematic sampling under non-response/missing observations. The new methods describe some ratio, product and regression-type estimators of the population mean when the information on a single auxiliary variable is utilized. A general class of estimators of the population mean has also been suggested. The theoretical results have been demonstrated by conducting an empirical study.

II. Sampling Design under Non-Response

Suppose a population consists of N units and the units are serially numbered from 1 to N . Let a sample of size n be

[1] Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, India, mk15@bhu.ac.in
[2] Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, India, anilstats0206@gmail.com
[3] Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, India, raybkray@gmail.com

[*]Corresponding Author

selected from the population such that $N = nk$ (k being the sampling interval and assumed to be an integer). First, we select a number $i(i \le k)$ randomly and choose the unit from the population corresponding to number i. Now, we select every k^{th} unit from the population. Thus, there would be k possible samples of same size n. Let Y and X be the study and auxiliary variables with respective population mean $\overline{Y}$ and $\overline{X}$. Let y_{ij} and $x_{ij}\left(i = 1,2,...,k;\ j = 1,2,...,n\right)$ be the observations on the j^{th} unit in the i^{th} systematic sample for the study and auxiliary variables respectively. Suppose the study variable is subject to the non-response and auxiliary variable is free from the non-response. Out of n units, let there be n_{i1} respondent units and n_{i2} non-respondent units for the study variable Y in the i^{th} systematic sample. To cope with the situation of non-response, we select a sub-sample of h_{i2} units from n_{i2} non-respondent units following the procedure of systematic sampling such that $n_{i2} = h_{i2}l$ (l being the sub-sampling interval and assumed to be an integer) and collect the information from all h_{i2} units by interview method. Thus, the estimator of the population mean $\overline{Y}$ based on $\left(n_{i1} + h_{i2}\right)$ units is given by

$$\overline{y}_{sys}^{*} = \overline{y}_{i.}^{*} = \frac{n_{i1}\overline{y}_{ni1} + n_{i2}\overline{y}_{hi2}}{n} \quad ;\ i = 1,2,...,k \tag{1}$$

where $\overline{y}_{ni1}$ and $\overline{y}_{hi2}$ are the means based on n_{i1} respondent units and h_{i2} non-respondent units respectively under study variable.

Now, the estimator of population mean $\overline{X}$ based on n units is given as

$$\overline{x}_{sys} = \overline{x}_{i.} = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \quad ;\ i = 1,2,...,k \tag{2}$$

The estimators $\overline{y}_{sys}^{*}$ and $\overline{x}_{sys}$ are respectively unbiased estimators of the population means $\overline{Y}$ and $\overline{X}$ because

$$E\left(\overline{y}_{sys}^{*}\right) = E\left(\overline{y}_{i.}^{*}\right) = E\left[E\left(\overline{y}_{i.}^{*} / n_{i1}, n_{i2}\right)\right] = E\left[\frac{n_{i1}\overline{y}_{ni1} + n_{i2}\overline{y}_{ni2}}{n}\right]$$

$$\Rightarrow E\left(\overline{y}_{sys}^{*}\right) = E\left[\overline{y}_{i.}\right] = \overline{Y} \tag{3}$$

and $E\left(\overline{x}_{sys}\right) = E\left(\overline{x}_{i.}\right) = \overline{X}$ (4)

where $\overline{y}_{ni2}$ and $\overline{y}_{i.}$ are the means based on n_{i2} non-respondent units and entire n units respectively under study variable.

The variances of the estimators $\overline{y}_{sys}^{*}$ and $\overline{x}_{sys}$ are respectively given by

$$V\left(\overline{y}_{sys}^{*}\right) = \frac{N-1}{nN}\left[1 + (n-1)\rho_Y\right]S_Y^2 + \frac{1}{n}\left[\begin{array}{l} W_2\left\{l + \left\{\frac{(NW_2 - 1)(n-1)}{N-1} + 1\right\} - (1+l)\rho_{Y2}\right\} \\ -\frac{l\left(1-\rho_{Y_2}\right)}{n} \end{array}\right]S_{Y2}^2 \tag{5}$$

$$V\left(\overline{x}_{sys}\right) = \frac{N-1}{nN}\left[1 + (n-1)\rho_X\right]S_X^2 \tag{6}$$

where S_Y^2 and S_X^2 are respectively the population mean squares under study and auxiliary variables. S_{Y2}^2 is the population mean square of the non-response group under study variable. ρ_Y and ρ_X are the correlation coefficients between a pair of units within the systematic sample for the study and auxiliary variables respectively. ρ_{Y2} is the correlation coefficient between a pair of units in the non-response group within the systematic sample for the study variable. W_2 is the non-response rate in the population.

The usual ratio, product and regression estimators of the population mean $\overline{Y}$ under non-response are respectively given as

$$\overline{y}_{R}^{*} = \frac{\overline{y}_{sys}^{*}}{\overline{x}_{sys}}\overline{X} \tag{7}$$

$$\overline{y}_{P}^{*} = \frac{\overline{y}_{sys}^{*}}{\overline{X}}\overline{x}_{sys} \tag{8}$$

$$\overline{y}_{lr}^{*} = \overline{y}_{sys}^{*} + b^{*}\left(\overline{X} - \overline{x}_{sys}\right) \tag{9}$$

where $b^{*} = \frac{s_{xyi}^{*}}{s_{xi}^{2}}$,

$s_{xyi}^{*} = \frac{1}{n-1}\left(\sum_{j \in s_1} x_{ij}y_{ij} + h_{i2}\sum_{j \in s_{2hi2}} x_{ij}y_{ij} - n\overline{x}_{sys}\overline{y}_{sys}^{*}\right)$ and

$s_{xi}^{2} = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{ij} - \overline{x}_{sys}\right)^2$. s_1 and s_{2hi2} are the sets of n_{i1} respondent units and h_{i2} non-respondent units respectively.

The expressions for the bias and mean square error (MSE) of the estimators $\overline{y}_R^*$, $\overline{y}_P^*$ and $\overline{y}_{lr}^*$ are respectively given by

$$Bias\left(\overline{y}_R^*\right) = \frac{N-1}{nN}\overline{Y}\left[1+(n-1)\rho_X\right]\left(1-K\rho^*\right)C_X^2 \tag{10}$$

$$MSE\left(\overline{y}_R^*\right) = \frac{N-1}{nN}\overline{Y}^2\left\{1+(n-1)\rho_X\right\}\left[\rho^{*2}C_Y^2+\left(1-2K\rho^*\right)C_X^2\right] + \frac{1}{n}\left[W_2\left\{l+\left\{\frac{(NW_2-1)(n-1)}{N-1}+1\right\}-(1+l)\rho_{Y_2}\right\}-\frac{l\left(1-\rho_{Y_2}\right)}{n}\right]S_{Y_2}^2 \tag{11}$$

$$Bias\left(\overline{y}_P^*\right) = E\left(\overline{y}_P^*-\overline{Y}\right) = \frac{N-1}{nN}\overline{Y}\left\{1+(n-1)\rho_X\right\}K\rho^*C_X^2 \tag{12}$$

$$MSE\left(\overline{y}_P^*\right) = \frac{N-1}{nN}\overline{Y}^2\left\{1+(n-1)\rho_X\right\}\left[\rho^{*2}C_Y^2+\left(1+2K\rho^*\right)C_X^2\right] + \frac{1}{n}\left[W_2\left\{l+\left\{\frac{(NW_2-1)(n-1)}{N-1}+1\right\}-(1+l)\rho_{Y_2}\right\}-\frac{l\left(1-\rho_{Y_2}\right)}{n}\right]S_{Y_2}^2 \tag{13}$$

$$MSE\left(\overline{y}_{lr}^*\right) = \frac{N-1}{nN}\overline{Y}^2\left\{1+(n-1)\rho_X\right\}\left[\rho^{*2}C_Y^2+\left(\frac{S_{XY}}{S_X^2}\frac{\overline{X}}{\overline{Y}}-2K\rho^*\right)\frac{S_{XY}}{S_X^2}\frac{\overline{X}}{\overline{Y}}C_X^2\right] + \frac{1}{n}\left[W_2\left\{l+\left\{\frac{(NW_2-1)(n-1)}{N-1}+1\right\}-(1+l)\rho_{Y_2}\right\}-\frac{l\left(1-\rho_{Y_2}\right)}{n}\right]S_{Y_2}^2 \tag{14}$$

where $C_X = \frac{S_X}{\overline{X}}$, $C_Y = \frac{S_Y}{\overline{Y}}$, $\rho^* = \frac{\left[1+(n-1)\rho_Y\right]^{1/2}}{\left[1+(n-1)\rho_X\right]^{1/2}}$, $K = \rho\frac{C_Y}{C_X}$, $S_{XY} = \rho S_X S_Y$. ρ is the population correlation coefficient between Y and X.

III. Suggested Estimators under New Estimation Strategy

We now consider another strategy of estimating the population mean $\overline{Y}$. Let there be $(n-p-q)$ completely available observations of Y and X, namely (y'_{i1}, x'_{i1}), (y'_{i2}, x'_{i2}), ..., $\left(y'_{i(n-p-q)}, x'_{i(n-p-q)}\right)$ on selected units in the i^{th} systematic sample of size n $(0 \le p, q \le n; 0 \le p+q \le n)$. Suppose there are p available observations of X, namely x''_{i1}, x''_{i2}, ..., x''_{ip} on the other units of the i^{th} systematic sample but the corresponding observations of Y are missing. Similarly, there are q available observations of Y, namely y''_{i1}, y''_{i2}, ..., y''_{iq} on the other units of the i^{th} systematic sample but the corresponding observations of X are missing. The quantities p and q are assumed to be random. Moreover, we have

$\overline{x}' = \frac{1}{n-p-q}\sum_j x'_{ij}$: Sample mean of X based on $(n-p-q)$ observations

$\overline{y}' = \frac{1}{n-p-q}\sum_j y'_{ij}$: Sample mean of Y based on $(n-p-q)$ observations

$\overline{x}'' = \frac{1}{p}\sum_j x''_{ij}$: Sample mean of X based on p observations

$\overline{y}'' = \frac{1}{q}\sum_j y''_{ij}$: Sample mean of Y based on q observations

$$\overline{x}_C = \frac{(n-p-q)\overline{x}' + p\overline{x}''}{n-q} \quad \text{and} \quad \overline{y}_C = \frac{(n-p-q)\overline{y}' + q\overline{y}''}{n-p}$$

Also, we have

$$V(\overline{x}') = \frac{N-1}{(n-p-q)N}\left\{1+(n-p-q-1)\rho_X\right\}S_X^2,$$

$$V(\overline{y}') = \frac{N-1}{(n-p-q)N}\left\{1+(n-p-q-1)\rho_Y\right\}S_Y^2,$$

$$V(\overline{x}'') = \frac{N-1}{pN}\left\{1+(p-1)\rho_X\right\}S_X^2$$

$$V(\overline{y}'') = \frac{N-1}{qN}\left\{1+(q-1)\rho_Y\right\}S_Y^2, \quad V\left(\overline{x}_C\right) = \frac{N-1}{N}\sigma_q^{*2}S_X^2,$$

$$V\left(\overline{y}_C\right) = \frac{N-1}{N}\sigma_p^{*2}S_Y^2 \quad \text{and} \quad Cov\left(\overline{y}_C, \overline{x}_C\right) = \frac{N-1}{N}\sigma_p^*\sigma_q^*\rho S_Y S_X$$

where

$$\sigma_p^{*2} = \frac{1}{(n-p)^2}\left[\phi_1+\phi_2+2\rho_Y\sqrt{\phi_1\phi_2}\right],$$

$\phi_1 = (n-p-q)\left\{1+(n-p-q-1)\rho_Y\right\}$, $\phi_2 = q\left\{1+(q-1)\rho_Y\right\}$,

$$\sigma_q^{*2} = \frac{1}{(n-q)^2}\left\{\psi_1+\psi_2+2\rho_X\sqrt{\psi_1\psi_2}\right\},$$

$\psi_1 = (n-p-q)\left\{1+(n-p-q-1)\rho_X\right\}$

and $\psi_2 = p\left\{1+(p-1)\rho_X\right\}$.

A. *Suggested Ratio, Product and Regression-Type Estimators*

We now define some ratio, product and regression-type estimators of the population mean $\overline{Y}$ in systematic sampling under non-response/missing observations as

$$\bar{y}'_R = \frac{\bar{y}_C}{\bar{x}_C}\bar{X} \tag{15}$$

$$\bar{y}'_P = \frac{\bar{y}_C}{\bar{X}}\bar{x}_C \tag{16}$$

$$\bar{y}'_{lr} = \bar{y}_C + b'\left(\bar{X} - \bar{x}_C\right) \tag{17}$$

where $b'\left(=\frac{s_{xy}}{s_x^2}\right)$ is an estimator of the population regression coefficient of Y on X i.e. $\beta = \frac{S_{XY}}{S_X^2}$. s_{xy} and s_x^2 are the unbiased estimators of the population parameters S_{XY} and S_X^2 based on $(n-p-q)$ and p units respectively.

In order to obtain the biases and mean square errors of the estimators defined in equations (15), (16) and (17), we use the theory of large sample approximations. Let us assume

$\bar{y}_C = \bar{Y}(1+e'_0)$, $\bar{x}_C = \bar{X}(1+e'_1)$, $s_{xy} = S_{XY}(1+e'_2)$, $s_x^2 = S_X^2(1+e'_3)$

such that $E(e'_0) = E(e'_1) = E(e'_2) = E(e'_3) = 0$,

$$E\left(e_0'^2\right) = \frac{V\left(\bar{y}_C\right)}{\bar{Y}^2} = \frac{N-1}{N}\frac{\sigma_p^{*2}}{\bar{Y}^2}S_Y^2 = \frac{N-1}{N}\sigma_p^{*2}C_Y^2,$$

$$E\left(e_1'^2\right) = \frac{V\left(\bar{x}_C\right)}{\bar{X}^2} = \frac{N-1}{N}\frac{\sigma_q^{*2}}{\bar{X}^2}S_X^2 = \frac{N-1}{N}\sigma_q^{*2}C_X^2 \text{ and}$$

$$E\left(e'_0e'_1\right) = \frac{Cov\left(\bar{y}_C, \bar{x}_C\right)}{\bar{Y}\bar{X}} = \frac{N-1}{N}\frac{\sigma_p^*}{\bar{X}}\frac{\sigma_q^*}{\bar{Y}}\rho S_X S_Y = \frac{N-1}{N}\sigma_p^*\sigma_q^*\rho C_X C_Y$$

.

Under the assumptions of large sample approximations, the expressions for the bias and MSE of the suggested estimators $\bar{y}'_R$, $\bar{y}'_P$ and $\bar{y}'_{lr}$ in terms of e'_0 , e'_1 , e'_2 and e'_3 up to the first order of approximation are respectively given as

$$Bias\left(\bar{y}'_R\right) = E\left(\bar{y}'_R - \bar{Y}\right) = \bar{Y}\left[E\left(e_1'^2\right) - E\left(e'_0e'_1\right)\right] \tag{18}$$

$$Bias\left(\bar{y}'_P\right) = E\left(\bar{y}'_P - \bar{Y}\right) = \bar{Y}E\left[e'_0e'_1\right] \tag{19}$$

$$MSE\left(\bar{y}'_R\right) = \bar{Y}^2\left[E\left(e_0'^2\right) + E\left(e_1'^2\right) - 2E\left(e'_0e'_1\right)\right] \tag{20}$$

$$MSE\left(\bar{y}'_P\right) = \bar{Y}^2\left[E\left(e_0'^2\right) + E\left(e_1'^2\right) + 2E\left(e'_0e'_1\right)\right] \tag{21}$$

$$MSE\left(\bar{y}'_{lr}\right) = \bar{Y}^2E\left(e_0'^2\right) + \frac{S_{XY}^2}{S_X^4}\bar{X}^2E\left(e_1'^2\right) - 2\frac{S_{XY}}{S_X^2}\bar{X}\bar{Y}E\left(e'_0e'_1\right) \tag{22}$$

Substituting the values of $E\left(e_0'^2\right)$, $E\left(e_1'^2\right)$ and $E\left(e'_0e'_1\right)$ in the equations (18), (19), (20), (21) and (22), we get the expressions for the bias and MSE of the suggested ratio, product and regression-type estimators as follows:

$$Bias\left(\bar{y}'_R\right) = \frac{N-1}{N}\bar{Y}\sigma_q^*C_X\left[\sigma_q^*C_X - \sigma_p^*\rho C_Y\right] \tag{23}$$

$$Bias\left(\bar{y}'_P\right) = \frac{N-1}{N}\bar{Y}\sigma_p^*\sigma_q^*\rho C_X C_Y \tag{24}$$

$$MSE\left(\bar{y}'_R\right) = \frac{N-1}{N}\bar{Y}^2\left[\sigma_p^{*2}C_Y^2 + \sigma_q^{*2}C_X^2 - 2\sigma_p^*\sigma_q^*\rho C_X C_Y\right] \tag{25}$$

$$MSE\left(\bar{y}'_P\right) = \frac{N-1}{N}\bar{Y}^2\left[\sigma_p^{*2}C_Y^2 + \sigma_q^{*2}C_X^2 + 2\sigma_p^*\sigma_q^*\rho C_X C_Y\right] \tag{26}$$

$$MSE\left(\bar{y}'_{lr}\right) = \frac{N-1}{N}S_Y^2\left[\sigma_p^{*2} + \rho^2\sigma_q^{*2} - 2\rho^2\sigma_p^*\sigma_q^*\right] \tag{27}$$

B. Suggested Class of Estimators

Following Gupta and Shabbir (2008), we now propose a general class of estimators for estimating the population mean $\bar{Y}$ under systematic sampling in the presence of non-response/missing observations as follows:

$$\bar{y}_t^* = \left[\alpha_1\bar{y}_C + \alpha_2\left(\bar{X} - \bar{x}_C\right)\right]\left(\frac{\eta\bar{X}+\lambda}{\eta\bar{x}_C+\lambda}\right) \tag{28}$$

where α_1 and α_2 are the suitably chosen constants, $\eta(\neq 0)$ and λ are either constants or the functions of known parameters of the auxiliary variable X .

Expressing the equation (28) in terms of e'_0 and e'_1 , we get

$$\bar{y}_t^* = \left[\alpha_1\bar{Y}\left(1+e'_0\right) + \alpha_2\left\{\bar{X} - \bar{X}\left(1+e'_1\right)\right\}\right]\left[\frac{\eta\bar{X}+\lambda}{\eta\bar{X}\left(1+e'_1\right)+\lambda}\right] \tag{29}$$

Solving the equation (29) and then ignoring the terms having powers of e'_0 and e'_1 greater than two, we have

$$\bar{y}_t^* - \bar{Y} = \left[\bar{Y}\left(\alpha_1 - 1\right) + \alpha_1\bar{Y}\left\{e'_0 - \tau e'_1 - \tau e'_0e'_1 + \tau^2e_1'^2\right\} - \alpha_2\bar{X}\left(e'_1 - \tau e_1'^2\right)\right] \tag{30}$$

where $\tau = \frac{\eta\bar{X}}{\eta\bar{X}+\lambda}$.

Taking expectation on both sides of equation (30), we get

$$E\left(\overline{y}_t^* - \overline{Y}\right) = \overline{Y}\left(\alpha_1 - 1\right) + \alpha_1 \overline{Y}\left\{\tau^2 E\left(e_1'^2\right) - \tau E\left(e_0' e_1'\right)\right\} + \alpha_2 \overline{X} \tau E\left(e_1'^2\right) \tag{31}$$

Substituting the values of $E\left(e_1'^2\right)$ and $E\left(e_0' e_1'\right)$ into equation (31), the expression for the bias of $\overline{y}_t^*$ up to the first order of approximation is given as

$$Bias\left(\overline{y}_t^*\right) = \left(\alpha_1 - 1\right)\overline{Y} + \frac{N-1}{N}\left[\alpha_1 \overline{Y}\left(\tau^2 \sigma_q^{*2} C_X^2 - \tau \sigma_p^* \sigma_q^* \rho C_X C_Y\right) + \alpha_2 \overline{X} \tau \sigma_q^{*2} C_X^2\right] \tag{32}$$

Squaring both sides of equation (30) and then taking expectation on neglecting the terms having powers of e_0' and e_1' higher than two, we get

$$E\left(\overline{y}_t^* - \overline{Y}\right)^2 = \left(\alpha_1 - 1\right)^2 \overline{Y}^2 + \alpha_1^2 \overline{Y}^2 \left\{E\left(e_0'^2\right) + \tau^2 E\left(e_1'^2\right) - 2\tau E\left(e_0' e_1'\right)\right\} + \alpha_2^2 \overline{X}^2 E\left(e_1'^2\right)$$
$$-2\alpha_1 \alpha_2 \overline{X}\,\overline{Y}\left\{E\left(e_0' e_1'\right) - \tau E\left(e_1'^2\right)\right\} \tag{33}$$

Substituting the values of $E\left(e_0'^2\right)$, $E\left(e_1'^2\right)$ and $E\left(e_0' e_1'\right)$ into equation (33), the expression for the MSE of $\overline{y}_t^*$ up to the first order of approximation is given by

$$MSE\left(\overline{y}_t^*\right) = \left(\alpha_1 - 1\right)^2 \overline{Y}^2 + \frac{N-1}{N}\begin{bmatrix}\alpha_1^2 \overline{Y}^2\left(\sigma_p^{*2} C_Y^2 - \tau^2 \sigma_q^{*2} C_X^2 - 2\tau \sigma_p^* \sigma_q^* \rho C_X C_Y\right) + \alpha_2^2 \overline{X}^2 \sigma_q^{*2} C_X^2 \\ -2\alpha_1 \alpha_2 \overline{X}\,\overline{Y} \sigma_q^* C_X\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right)\end{bmatrix} \tag{34}$$

To obtain the optimum values of α_1 and α_2, we differentiate the equation (34) with respect to α_1 and α_2 respectively and equate the derivatives to zero. Thus, we have

$$\frac{\partial MSE\left(\overline{y}_t^*\right)}{\partial \alpha_1} = 2\left(\alpha_1 - 1\right)\overline{Y}^2 + \frac{N-1}{N}\begin{bmatrix}2\alpha_1 \overline{Y}^2\left(\sigma_p^{*2} C_Y^2 - \tau^2 \sigma_q^{*2} C_X^2 - 2\tau \sigma_p^* \sigma_q^* \rho C_X C_Y\right) \\ -2\alpha_2 \overline{X}\,\overline{Y} \sigma_q^* C_X\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right)\end{bmatrix} = 0$$

$$\Rightarrow \left(\alpha_1 - 1\right)\overline{Y}^2 + \frac{N-1}{N}\begin{bmatrix}\alpha_1 \overline{Y}^2\left(\sigma_p^{*2} C_Y^2 - \tau^2 \sigma_q^{*2} C_X^2 - 2\tau \sigma_p^* \sigma_q^* \rho C_X C_Y\right) \\ -\alpha_2 \overline{X}\,\overline{Y} \sigma_q^* C_X\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right)\end{bmatrix} = 0 \tag{35}$$

$$\frac{\partial MSE\left(\overline{y}_t^*\right)}{\partial \alpha_2} = \frac{N-1}{N}\left[2\alpha_2 \overline{X}^2 \sigma_q^{*2} C_X^2 - 2\alpha_1 \overline{X}\,\overline{Y} \sigma_q^* C_X\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right)\right] = 0$$

$$\Rightarrow \alpha_2 \overline{X}^2 \sigma_q^{*2} C_X^2 - \alpha_1 \overline{X}\,\overline{Y} \sigma_q^* C_X\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right) = 0 \tag{36}$$

Solving the equations (35) and (36), we can get the optimum values of α_1 and α_2 as

$$\alpha_{1(opt)} = \frac{1}{1 + \frac{N-1}{N}\left(1 - \rho^2\right)\sigma_p^{*2} C_Y^2} \tag{37}$$

$$\alpha_{2(opt)} = \frac{\alpha_{1(opt)} \overline{Y}\left(\sigma_p^* \rho C_Y - \tau \sigma_q^* C_X\right)}{\overline{X} \sigma_q^* C_X} \tag{38}$$

Substituting the optimum values of α_1 and α_2 from equations (37) and (38) into the equation (34), one can get the expression for the MSE of the optimum estimator $\overline{y}_{t(opt)}^*$ as

$$MSE\left(\overline{y}_{t(opt)}^*\right) = \frac{MSE\left(\overline{y}_{lr}'\right)}{1 + \frac{N-1}{N}\left(1 - \rho^2\right)\sigma_p^{*2} C_Y^2} \tag{39}$$

IV. Empirical Study

To verify the performance of the suggested estimators, we have illustrated the theoretical results obtained in the previous sections through a real data set. We have used the data considered by Shahzad *et al.* (2017). In this data set, volume of the timber and length of the timber are respectively considered as Y and X. The details are given below:

$N = 176$, $n = 16$, $\overline{Y} = 282.61$, $\overline{X} = 6.99$, $S_Y = 155.73$, $S_X = 2.95$, $\rho = 0.87$, $\rho_Y = -0.0019$, $\rho_X = -0.0014$, $p = 3,4,5,6,7$ and $q = 3,4,5$.

Table 1 shows the variance/MSE of the estimators $\overline{y}_C$, $\overline{y}_R'$, $\overline{y}_P'$, $\overline{y}_{lr}'$ and $\overline{y}_{t(opt)}^*$ for the different choices of p and q.

Table 1: Variance/MSE of $\overline{y}_C$, $\overline{y}_R'$, $\overline{y}_P'$, $\overline{y}_{lr}'$ and $\overline{y}_{t(opt)}^*$

p	q	Variance /MSE				
		$\overline{y}_C$	$\overline{y}_R'$	$\overline{y}_P'$	$\overline{y}_{lr}'$	$\overline{y}_{t(opt)}^*$
3	3	1825.9613	463.0848	5339.8826	443.8972	441.4437
4	3	1981.4108	516.6281	5599.8072	483.7848	480.8846
5	3	2165.0308	584.8895	5900.5096	536.4963	532.9840
4	4	1984.3021	503.3714	5800.8709	482.3880	479.4920
5	4	2167.7652	566.9534	6105.7238	529.7887	526.3159
6	4	2387.7059	649.7357	6463.3270	593.9632	589.6777
5	5	2169.1404	550.3539	6339.7404	527.3212	523.8625
6	5	2388.5	626.8	6701.9	584.4	580.

		303	825	918	278	2096
7	5	2656.2 116	729.0 259	7134.0 225	663.9 746	658. 6495

Table 2 represents the percentage relative efficiency (PRE) of the estimators $\overline{y}_C$, $\overline{y}'_R$, $\overline{y}'_P$, $\overline{y}'_{lr}$ and $\overline{y}^*_{t(opt)}$ with respect to the estimator $\overline{y}_C$ for the different choices of p and q.

Table 2: PRE of $\overline{y}_C$, $\overline{y}'_R$, $\overline{y}'_P$, $\overline{y}'_{lr}$ and $\overline{y}^*_{t(opt)}$

p	q	*PRE*				
		$\overline{y}_C$	$\overline{y}'_R$	$\overline{y}'_P$	$\overline{y}'_{lr}$	$\overline{y}^*_t$
3	3	100	394.3 039	34.19 48	411.3 478	413.6 340
4	3	100	383.5 275	35.38 36	409.5 645	412.0 346
5	3	100	370.1 606	36.69 23	403.5 500	406.2 093
4	4	100	394.2 024	34.20 70	411.3 498	413.8 342
5	4	100	382.3 533	35.50 38	409.1 755	411.8 753
6	4	100	367.4 888	36.94 24	401.9 956	404.9 171
5	5	100	394.1 356	34.21 50	411.3 509	414.0 667
6	5	100	381.0 173	35.63 91	408.6 955	411.6 668
7	5	100	364.3 508	37.23 30	400.0 472	403.2 815

Concluding Remarks

In this paper, some new methods of estimating the population mean have been discussed under systematic sampling in the presence of non-response/missing observations. Some ratio, product and regression-type estimators of the population mean have been suggested under the situation in which certain observations for some sampling units are missing. A general class of estimators of the population mean has also been suggested under the given situation. The fundamental properties of the suggested estimators have been discussed in detail. To examine the behaviour of the suggested estimators, an empirical study has been carried out through a real data set. The Table 1 and Table 2 reveal that a significant gain in the efficiency of the suggested estimators has been achieved. Though, the product-type estimator $\overline{y}'_P$ is less efficient than the estimator $\overline{y}_C$. Such a result could be seen because of the positive correlation between study and auxiliary variables. The estimator $\overline{y}'_P$ would perform well if both the variables are negatively correlated. From Table 1 and Table 2, it is also revealed that the optimum estimator of the proposed class performs very well as compared to the other estimators.

Acknowledgement

This work is supported by Incentive Grant, IoE, Banaras Hindu University.

References

Buckland, W. R. (1951). A review of the literature of systematic sampling. *Journal of the Royal Statistical Society: Series B (Methodological), 13*(2), 208–215.

Chaudhary, M. K., Malik, S., Singh, J., & Singh, R. (2012). A general family of estimators for estimating population mean in systematic sampling using auxiliary information in the presence of missing observations. *Journal of Rajasthan Statistical Association, 1*(2), 1–8.

Chaudhary, M. K., Malik, S., Singh, R., & Smarandache, F. (2013). Use of auxiliary information for estimating population mean in systematic sampling under non-response. *On improvement in estimating population parameter(s) using auxiliary information*. Education Publishing, 5–16.

Cochran, W. G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *The Annals of Mathematical Statistics, 17*(2), 164–177.

Finney, D. J. (1948). Random and systematic sampling in timber surveys. *Forestry, 22*(1), 64–99.

Grover, L. K., & Kaur, P. (2014). Exponential ratio type estimators of population mean under non-response. *Open Journal of Statistics, 4,* 97–110.

Gupta, S., & Shabbir, J. (2008). On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics, 35*(5), 559–566.

Hansen, M. H., & Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association, 41*(236), 517–529.

Hasel, A. A. (1942). Estimation of volume in timber stands by strip sampling. *The Annals of Mathematical Statistics, 13*(2), 179–206.

Khan, M., & Singh, R. (2015). Estimation of Population Mean in Chain Ratio-Type Estimator under Systematic Sampling. *Journal of Probability and Statistics*, *2015*(1), 1-5.

Kushwaha, K. S., & Singh, H. P. (1989). Class of almost unbiased ratio and product estimators in systematic sampling.

Journal of the Indian Society of Agricultural Statistics, 41(2), 193–205.

Lahiri, D. B. (1954). *Technical paper on some aspects of the development of the sample design* (Technical Paper No. 5). The National Sample Survey, Department of Economic Affairs, Ministry of Finance, Government of India.

Madow, L. H. (1946). Systematic sampling and its relation to other sampling designs. *Journal of the American Statistical Association, 41*(234), 204–217.

Madow, W. G., & Madow, L. H. (1944). On the theory of systematic sampling. *The Annals of Mathematical Statistics, 15*(1), 1–24.

Mishra, M., Singh, B. P., & Singh, R. (2018). Estimation of population mean using two auxiliary variables in systematic sampling. *Journal of Scientific Research*, *62*, 203–212.

Nair, K. R., & Bhargava, R. P. (1951). *Statistical sampling in timber surveys in India* (Indian Forest Leaflet No. 153). Forest Research Institute.

Shahzad, U., Hanif, M., Koyuncu, N., & Luengo, A. V. G. (2017). Estimation of population mean under systematic random sampling in absence and presence of non-response. *Journal of Statistics: Statistics and Actuarial Sciences, 1*, 23–39.

Singh, R., Malik, S., Chaudhary, M. K., Verma, H. K., & Adewara, A. A. (2012). A general family of ratio-type estimators in systematic sampling. *Journal of Reliability and Statistical Studies*, *5*(1), 73–82.

Singh, R., & Singh, H. P. (1998). Almost unbiased ratio and product type estimators in systematic sampling. *Qüestiió*, *22*(3), 403–416.

Sukhatme, P. V., Panse, V. G., & Sastry, K. V. R. (1958). Sampling technique for estimating the catch of sea fish in India. *Biometrics*, *14*(1), 78–96.

Toutenburg, H., & Srivastava, V. K. (1998). Estimation of ratio of population means in survey sampling when some observations are missing. *Metrika*, *48*, 177–187.

Yates, F. (1948). Systematic sampling. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences*, *241*(834), 345–377.

***