# ESTIMATATION OF POPULATION MEAN UNDER DIAGONAL SYSTEMATIC SAMPLING SCHEME BY USING IMPUTATION METHODS

**Abstract**

In survey sampling, the use of auxiliary information to enhance estimators of population parameters under simple random sampling stratified random sampling and systematic sampling has been widely discussed. Similarly, some existing estimators were modified using regression imputation approach to obtain two imputation schemes and estimators that impute the responses non-respondents thereby eliminating difficulties in data presentation, compilation. The theoretical properties (estimators, biases and mean squared errors) of the proposed imputation scheme were derived so as to assess their robustness and efficiency. The theoretical findings were supported by simulation studies on population generated using four distributions namely; Beta, Gamma, Exponential and Uniform distributions. The averages of biases, MSEs and PREs of the estimators in comparison to the existing estimators were computed from the simulated data and the results showed that on average, the estimators of the proposed imputation scheme have minimum biases, minimum MSEs and higher PREs compared to the traditional unbiased estimators. These results imply that the estimators of the proposed schemes are more efficient and robust than the conventional unbiased estimators.

**Keywords:** Estimator, Efficiency, Robustness, Distributions, Systematic Sampling

## 1.0 Introduction

The use of auxiliary information to enhance estimators of population parameters under simple random sampling stratified random sampling and systematic sampling has been widely discussed. Cochran (1940) used auxiliary information at estimation stage and suggested a ratio estimator. The ratio estimator is more efficient when study and auxiliary variates are positively correlated and regression line passes through the origin. In case of negative correlation, Robson (1957) developed product method of estimation that provides a product estimator which is more efficient than the simple mean estimator. Hansen et al. (1946) developed a combined ratio estimator in stratified sampling. In systematic sampling, Swain (1964) defined a ratio estimator whereas Shukla (1971) suggested a product estimator. Many authors including Kushwaha and Singh (1989), Singh and Singh (1998) and Singh and Solanki (2012) discussed various estimators of population mean. Singh et al. (2011) suggested a general family of estimators for estimating population mean in systematic sampling using auxiliary information in the presence of missing observations. Singh and Jatwa (2012) suggested a class of exponential-type estimators in systematic sampling. Singh *et al*. (2011) studied some modified ratio and product estimators for population mean in systematic sampling. Singh (1967) used information on population means of two auxiliary variates and developed a ratio-cum-product estimator in simple random sampling. Ayed et al. (2023) Suggested that the efficiencies of the estimators of the population parameters of the study variable can be increased by the use of auxiliary information related to auxiliary variable x, which is highly correlated with the study variable y. Auxiliary information may be efficiently utilized either at planning stage or at design stage to arrive at an improved estimator compared to those estimators, not utilizing auxiliary information. A simple technique of utilizing the known information of the population parameters of the auxiliary variables is through ratio, product, and regression method of estimations using different probability sampling designs such as simple random sampling, stratified random sampling, cluster sampling, double sampling. In this study we consider auxiliary information under the frame work of systematic sampling.

Systematic sampling has gotten the attention of survey statisticians due to its simplicity of use (Azeem and Khan, 2021). Systematic sampling is even simpler than simple random sampling as only the first units (or the first few units) are selected randomly from the population. The remaining units are obtained according to a pre-defined rule. First introduced by Madow and Madow (1944), many versions of systematic sampling have been developed by the researchers for use with different real-life situations. Madow and Madow (1944) introduced the novel idea of selecting the units the population according to a pre-defined pattern, called systematic sampling. Madow and Madow (1944) proposed method was only applicable to those circumstances where the size of the finite population is a constant multiple of the required sample size, thus limiting its usability. To overcome this drawback, Lahiri (1951) introduced a new method called circular systematic sampling design. Later, Chang and Huang (2000) introduced a new modification of systematic random sampling which they called remainder systematic sampling which is also applicable in situations in which the size of a finite population is not a multiple of the sample size. Apart from its simplicity, systematic sampling provides estimators which are more efficient than simple random sampling or stratified random Sampling for certain types of population; see Cochran (1946), Gautschi (1957)and Hajeck (1959). Notable among them are Singh et al. (2011), Singh et al. (2012), Tailor et al. (2013), Khan et al. (2013), Subramani (2013), Verma et al. (2014), Verma and Singh (2014), Khan and Rajesh (2015), Noor-ul-Amin et al (2017), Azeem and Khan (2021), Zahoor et al. (2022), Ayed and Khan (2023), Azeem (2023).

The diagonal systematic sampling procedure was first introduced by Subramani (2000) which was found to be more precise as compared to simple random sampling and linear systematic sampling methods in the presence of linear trend. Sampath and Varalakshmi (2008) introduced a new modified systematic sampling method called diagonal circular systematic sampling. Subramani (2009) introduced generalized version of diagonal systematic sampling method. Khan et al. (2014) suggested the conditions under which the Sampath and Varalakshmi (2008) sampling scheme is applicable. Khan et al. (2015) proposed a generalized version of systematic sampling, and it was shown that diagonal systematic sampling scheme is a special case of the new generalized sampling scheme. Imputation is a statistical method used to replace missing data with estimated values. Data obtained from sampling surveys often face the problem of non-response or missing values. These missing values create difficulty in analysis, processing and handling of data. The problem of non-response has been considered by many authors including Singh and Horn (2000), Singh and Deo (2003), Wang and Wang (2006), Kadilar and Cingi

(2008),Audu et.al. In this paper, we intendproposing imputation schemes that impute responses for non-respondents while using Subramani and Gupta (2014) estimators and to obtained new modified estimators for the schemes as well as their biases and mean square errors.

### 1.1 Literature review

Consider a finite population $U = U_1, U_{2,}U_{3,}.......U_N$ of size N units. A sample of size n is taken at random from the first $k$ units and every $k^{th}$ subsequent unit then, N = $nk$ where $n$ and $k$ are positive integers thus, there will be $k$ samples each of size $n$ and observe the study variable $y$ and auxiliary variate $x$ for each and every unit selected in the sample.

Let $(y_{ij,}x_{ij})$ for $i = 1,2......k,\ j = 1,2.....n$ indicate the value of $j^{th}$ unit in the $i^{th}$ sample and each value of $(y_{ij,}x_{ij})$ may be classify into two mutually exclusive classes, i.e. $G$ and $G^{'}$, where $G$ refers to the class of interest. Let

$$y_{ij} = \begin{cases} 1 & if\ \ y_{ij} \in G \\ 0 & if\ \ y_{ij} \in G^{'} \end{cases} \qquad (1.1)$$

$$x_{ij} = \begin{cases} 1 & if\ \ x_{ij} \in G \\ 0 & if\ \ x_{ij} \in G^{'} \end{cases} \qquad (1.2)$$

Then, the systematic sample means are defined as follows:

$\bar{y}_{st} = t_0 = 1/n\sum_{j=1}^{n} y_{ij}$ and $\bar{x}_{st} = t_0 = 1/n\sum_{j=1}^{n} y_{ij}$ unbiased estimators of the population means $\bar{Y} = 1/n\sum_{j=1}^{n} y_{ij}$ , and $\bar{X} = 1/n\sum_{j=1}^{n} x_{ij}$ of $y$ and $x$

To obtain estimators up to first order of approximation, using the following error terms:

$e_0 = \bar{y}_{sys} - \bar{Y} / \bar{X}, e_1 = \bar{x}_{sys} - \bar{X} / \bar{X}, e_2 = \bar{z}_{sys} - \bar{z} / \bar{z}$

Subramani and Gupta (2014) proposed the modified systematic sampling scheme which is explained by the following step.

1. Firstly arrange the N population units (labels) in a matrix with $k = k_1 + k_2$ columns and That is, the first $n_2k$ population units are arranged in row wise in the first $n_2$ rows with k elements each and the remaining $(n_1 - n_2)k_1$ population units are arranged row wise in the next $(n_1 - n_2)$ rows with k1 elements each as in the arrangement.
2. The first $k_1$ columns are assumed as Set 1 and the next $k_2$ columns are assumed are assumed at set 2
3. Select two random numbers, $i$ in between 1 and $k_1$ and j in between 1 and $k_2$ ,then select all the $n_1$ units in the $i^{th}$ column of Set 1 and all the $n_2$ units in the $j^{th}$ column of Set 2, which together give the sample of size $n$.
4. The step 3 leads to $k_1 \times k_2$ samples of size $n$

List 1: The arrangement of population unit of set-1 is given below.

| S.No | 1 | 2 | 3 | 4 | K |
|---|---|---|---|---|---|
| 1 | 1 | . | $i$ | . | $k$ |

| | | | | | |
|---|---|---|---|---|---|
| 2 | $k+1$ | . | $k+i$ | . | $k+k_1$ |
| 3 | $2k+1$ | . | $2k+i$ | . | $2k+k_1$ |
| 4 | . | . | . | . | . |
| 5 | . | . | . | . | . |
| 6 | . | . | . | . | . |
| 7 | $(n_2-1)k+1$<br>$n_2k+1$ | . | $(n_2-1)k+i$<br>$n_2k+i$ | . | $(n_2-1)k+k_1$<br>$n_2k+k_1$ |
| 8 | $n_2k+k_1+1$ | . | $n_2k+k_1+i$ | . | . |
| 9 | . | . | . | . | . |
| K | $n_2k_2+(n_1-1)k_1+1$ | . | $n_2k_2+(n_1-1)k_1+i$ | . | $n_2k_2+n_1k_1$ |

List 2 : The arrangement of population unit of set-2 is given below

| S.No | 1 | 2 | 3 | 4 | K |
|---|---|---|---|---|---|
| 1 | $k_1+1$ | . | $k_1+j$ | . | $2k$ |
| 2 | $k+k_1+1$ | . | $k+k_1+j$ | . | $3k$ |
| 3 | $2k+k_1+1$ | . | $2k+k_1+j$ | . | . |
| 4 | . | . | . | . | . |
| 5 | . | . | . | . | . |
| 6 | . | . | . | . | . |
| 7 | $(n_2-1)k_1+k_1+1$ | . | $(n_2-1)k_1+k_1+j$ | . | $n_2k$ |

$$\pi_i=\begin{cases}\dfrac{1}{k_1} & \text{if } i^{th} \text{ unit is from the set}-1\\[2ex] \dfrac{1}{k_2} & \text{if } i^{th} \text{ unit is from the set}-2\end{cases}$$

$$\pi_{ij}=\begin{cases}\dfrac{1}{k_1}, & \text{if } i^{th} \text{ and } j^{th} \text{ units are from the column of set}-I\\[2ex] \dfrac{1}{k_2}, & \text{if } i^{th} \text{ and } j^{th} \text{ units are from the column of set}-2\\[2ex] \dfrac{1}{k_1\times k_2}, & \text{if } i^{th} \text{ and } j^{th} \text{ units are from the column of sets}-1 \text{ and set}-2 \text{ respectively,}\\[2ex] 0, & \text{Otherwise.}\end{cases}$$

In general, for the given population size $N=n_1k_1+n_2k_2$, the selected generalized modified linear systematic samples (labels of the population units) for the random starts $i$ and $j$ are given below:

$$S_{ij}=\begin{cases}i, i+k, i+2k, ...., i+(n_2-1)k, i+n_2k, i+n_2k+k_1, ...., i+n_2k+(n_1-n_2-1)k_1,\\ j+k_1, j+k_1+k, ...., j+k_1+(n_2-1)k \quad (i=1,2,...,k_1 \text{ and } j=1,2,3,....,k_2)\end{cases} \tag{1.3}$$

### 1.2 Estimator of Population Mean In Diagonal Systematic Sapling and Its Properties

The generalized modified linear systematic sample mean based on the random starts $i$ and $j$ is show below as

$$\bar{y}_{gmlss}=\frac{1}{n}\left(\sum_{i=0}^{n_2-1}y_{i+kl}+\sum_{i=o}^{n_1-n_2-1}y_{i+n_2k+k_1l}+\sum_{j=0}^{n_2-1}y_{j+k_1+kl}\right)\quad for\ i=1,2,...k_1\ and\ j=1,2,3,....k_2 \tag{1.4}$$

Since the first order inclusion probabilities are not equal, the generalized modified linear systematic sample mean given above in (1.4) is not an unbiased estimator. The mean squared error of the GMLSS mean can be obtained from (1.5) as given below

$$MSE\left(\bar{y}_{gmlss}\right)=\frac{1}{k_1k_2}\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\left(\bar{y}_{ij}-\bar{Y}\right)^2 \tag{1.5}$$

### 2.0 Materials and Method

Let consider $(y,x)\ \varepsilon\ \psi^+$ be pair of the study associated auxiliary measured on the studied population. Also let $H$ denote the set of response with $n_1$ units, $H^c$ denotes the set of non-responses having $n-n_1$ units or missing units (out of $n$) and $S$ denotes the set of $n$ units sampled without replacement from the $N$ units in the population of interest. For each $i\ \varepsilon\ H$, the value of $y_i$ isobserved. Likewise, for unit $i\ \varepsilon\ H^c$, is missing due to non-response and obtainedusing different procedure of imputation.

### 2.1First Proposed Imputation Schemes and Its Estimator

Motivated by Auduet al. (2021), the regression type imputation scheme for population mean based on diagonal systematic sampling design is proposed as in (1.6)

$$y_{.i}=\begin{cases} y_i & if\ \ i\ \varepsilon\ H \\ \dfrac{\bar{y}_{gmls(r^*)}+\hat{\beta}\left(\bar{X}-\bar{x}_{gmls(r^*)}\right)}{M_1\bar{x}_{gmls(r^*)}+M_2}\left(M_1\bar{X}+M_2\right) & if\ \ i\ \varepsilon\ H^c \end{cases} \tag{1.6}$$

where,

$$\bar{y}_{gmlss(r^*)}=\frac{1}{r^*}\left(\sum_{i=0}^{r_1^*-1}y_{i+kl}+\sum_{i=o}^{r_1^*-r_2^*-1}y_{i+n_2k+k_1l}+\sum_{j=0}^{r_2^*-1}y_{j+k_1+kl}\right),\ \bar{x}_{gmlss(r^*)}=\frac{1}{r^*}\left(\sum_{i=0}^{r_1^*-1}x_{i+kl}+\sum_{i=o}^{r_1^*-r_2^*-1}x_{i+n_2k+k_1l}+\sum_{j=0}^{r_2^*-1}x_{j+k_1+kl}\right)$$

$$\hat{\beta}=s_{yx(r^*)}/s^2_{x(r^*)},\quad s_{yx(r^*)}=\left(r^*-1\right)^{-1}\sum_{i=1}^{r^*}\left(y_i-\bar{y}^*_{gmls(r^*)}\right)\left(x_i-\bar{x}^*_{gmls(r^*)}\right),\ s^2_{x(r^*)}=\left(r^*-1\right)^{-1}\sum_{i=1}^{r^*}\left(x_i-\bar{x}^*_{gmls(r^*)}\right)^2$$

$$r^*=r_1^*+r_2^*$$

The estimator of the first proposed imputation can be obtained as in (1.7).

$$t_1 = \frac{1}{n}\left(\sum_{i\in H}^{r} y_i + \sum_{i\in H^c}^{n-r} \frac{\bar{y}_{gmls(r^*)} + \hat{\beta}\left(\bar{X} - \bar{x}_{gmls(r^*)}\right)}{M_1\bar{x}_{gmls(r^*)} + M_2}\right)\left(M_1\bar{X} + M_2\right) \tag{1.7}$$

Simplify (1.7), (1.8) is obtained.

$$t_1 = \frac{r^*}{n}\bar{y}_{gmls(r^*)} + \left(1 - \frac{r^*}{n}\right)\frac{\left(\bar{y}_{gmls(r^*)} + \hat{\beta}\left(\bar{X} - \bar{x}_{gmls(r^*)}\right)\right)}{M_1\bar{x}_{gmls(r^*)} + M_2}\left(M_1\bar{X} + M_2\right) \tag{1.8}$$

Express (1.7) in terms of error term defined in (1.6), (1.9) is obtained.

$$t_1 = \frac{r^*}{n}\left(\bar{Y} + \bar{Y}e_0\right) + \left(1 - \frac{r^*}{n}\right)\frac{\left(\bar{Y} + \bar{Y}e_0 + \hat{\beta}\left(\bar{X} - \left(\bar{X} + \bar{X}e_1\right)\right)\right)}{M_1\left(\bar{X} + \bar{X}e_1\right) + M_2}\left(M_1\bar{X} + M_2\right) \tag{1.9}$$

$$t_1 = \frac{r^*}{n}\left(\bar{Y} + \bar{Y}e_0\right) + \left(1 - \frac{r^*}{n}\right)\left(\bar{Y} + \bar{Y}e_0 - \hat{\beta}\bar{X}e_1\right)\left(1 + \lambda e_1\right)^{-1} \tag{1.10}$$

where $\lambda = \dfrac{M_1\bar{X}}{M_1\bar{X} + M_2}$

Simplify (1.10) up to first order approximation, (1.11) is obtained

$$t_1 = \bar{Y} + \bar{Y}e_0 + \left(1 - \frac{r^*}{n}\right)\left(-\hat{\beta}\bar{X}e_1 - \lambda\bar{Y}e_1 - \lambda\bar{Y}e_0e_1 + \lambda\hat{\beta}\bar{X}e_1^2 + \bar{Y}\lambda^2e_1^2\right) \tag{1.11}$$

Subtract $\bar{Y}$ from both side of (1.11), (1.12) is obtained.

$$t_1 - \bar{Y} = \bar{Y}e_0 - \left(1 - \frac{r^*}{n}\right)\left(\hat{\beta}\bar{X} + \lambda\bar{Y}\right)e_1 + \left(1 - \frac{r^*}{n}\right)\left(\lambda\hat{\beta}\bar{X} + \bar{Y}\lambda^2\right)e_1^2 - \left(1 - \frac{r^*}{n}\right)\lambda\bar{Y}e_0e_1 \tag{1.12}$$

To obtain the bias of $t_1$, take expectation (1.12).

Since $E\left(e_0\right) = E\left(e_1\right) = 0,\ E\left(e_0^2\right) = \left(\frac{1}{r} - \frac{1}{N}\right)\frac{S_y^2}{\bar{Y}^2},\ E\left(e_0e_1\right) = \left(\frac{1}{r} - \frac{1}{N}\right)\frac{S_{xy}}{\bar{Y}\,\bar{X}}$, the bias $t_1$ is obtain as in (1.13)

$$Bias\left(t_1\right) = \left(1 - \frac{r^*}{n}\right)\left(\frac{1}{r} - \frac{1}{N}\right)\left[\left(\lambda\hat{\beta}\bar{X} + \bar{Y}\lambda^2\right)C^2x - \lambda\bar{Y}\rho yxCyCx\right] \tag{1.13}$$

To obtain the MSE of $t_1$, square (1.13) and take expectation of the result, the MSE of $t_1$ is obtained as in (1.14).

$$MSE\left(t_1\right) = \left(\frac{1}{r} - \frac{1}{N}\right)\left(\bar{Y}^2C_y^2 + \left[\left(1 - \frac{r^*}{n}\right)^2\left(\hat{\beta}\bar{X} + \lambda\bar{Y}\right)^2C_x^2\right] - 2\bar{Y}\left(1 - \frac{r^*}{n}\right)\left(\hat{\beta}\bar{X} + \lambda\bar{Y}\right)\rho_{xy}C_xC_y\right) \tag{1.14}$$

**2.2SecondProposed Imputation Schemes and Its Estimator**

Motivated by Audu and Singh (2021), the regression type imputation scheme for population mean based on diagonal systematic sampling design is proposed as in (1.15)

$$y_{.i}=\begin{cases} y_i & if\ \ i\ \varepsilon\ H \\ \dfrac{\bar{y}_{gmls(r^*)}+\hat{\beta}\left(\bar{X}-\bar{x}_{gmls(r^*)}\right)}{M_1\bar{x}_{gmls(r^*)}+M_2}\left(M_1\bar{X}+M_2\right)\exp\left(\dfrac{\bar{X}-\bar{x}^*_{gmsl(r^*)}}{\bar{X}+\bar{x}^*_{gmls(r^*)}}\right) & if\ \ i\ \varepsilon\ H^c \end{cases} \tag{1.15}$$

The estimators of the second proposed imputation scheme can be expressed as in (1.16)

$$t_2=\frac{1}{n}\left(\sum_{i\in H}^{r}y_i+\sum_{i\in H^c}^{n-r}\frac{\bar{y}_{gmls(r^*)}+\hat{\beta}\left(\bar{X}-\bar{x}_{gmls(r^*)}\right)}{M_1\bar{x}_{gmls(r^*)}+M_2}\right)\left(M_1\bar{X}+M_2\right)\exp\left[\frac{\bar{X}-\bar{x}_{gmls(r^*)}}{\bar{X}+\bar{x}_{gmls(r^*)}}\right] \tag{1.16}$$

Simplify (1.16) and (1.17) is obtained.

$$t_2=\frac{r^*}{n}\bar{y}_{gmls(r^*)}+\left(1-\frac{r}{n}\right)\frac{\left(\bar{y}_{gmls(r^*)}+\hat{\beta}\left(\bar{X}-\bar{x}_{gmls(r^*)}\right)\right)}{M_1\bar{x}_{gmls(r^*)}+M_2}\left(M_1\bar{X}+M_2\right)\exp\left[\frac{\bar{X}-\bar{x}_{gmls(r^*)}}{\bar{X}+\bar{x}_{gmls(r^*)}}\right] \tag{1.17}$$

Express (1.17) in terms of error terms defined in (1.15), (1.18) is obtained.

$$t_2=\frac{r^*}{n}\left(\bar{Y}+\bar{Y}e_0\right)+\left(1-\frac{r}{n}\right)\frac{\left(\left(\bar{Y}+\bar{Y}e_0\right)+\hat{\beta}\left(\bar{X}-\left(\bar{X}+\bar{X}e_1\right)\right)\right)}{M_1\left(\bar{X}+\bar{X}e_1\right)+M_2}\left(M_1\bar{X}+M_2\right)\exp\left[\frac{\bar{X}-\left(\bar{X}+\bar{X}e_1\right)}{\bar{X}+\left(\bar{X}+\bar{X}e_1\right)}\right] \tag{1.18}$$

Simplify (1.18) up to first order approximation, (1.19) is obtained.

$$t_2=\frac{r^*}{n}\bar{Y}+\frac{r^*}{n}\bar{Y}e_0+\left(1-\frac{r^*}{n}\right)\left(\bar{Y}+\bar{Y}e_0-\hat{\beta}\bar{X}e_1\right)\left(1-\lambda e_1+\lambda e_1^2\right)\left[1-\frac{e_1}{2}+\frac{3e_1^2}{8}\right] \tag{1.19}$$

Subtract $\bar{Y}$ up to first order approximation (1.19), (1.20) is obtained.

$$\begin{aligned} t_2-\bar{Y}&=\bar{Y}e_0+\left(1-\frac{r^*}{n}\right)\left(-\hat{\beta}\bar{X}-\lambda\bar{Y}-\frac{\bar{Y}}{2}\right)e_1 \\ &+\left(1-\frac{r^*}{n}\right)\left(\lambda\hat{\beta}\bar{X}+\lambda^2\bar{Y}+\frac{\hat{\beta}\bar{X}}{2}+\frac{\lambda\bar{Y}}{2}+\frac{3\bar{Y}}{8}\right)e_1^2+\left(1-\frac{r^*}{n}\right)\left(-\lambda\bar{Y}-\bar{Y}\right)e_0e_1 \end{aligned} \tag{1.20}$$

To obtain the bias of $t_2$, take expectation (1.20), the bias $t_2$ is obtained as in (1.21).

$$Bias\left(t_2\right)=\left(1-\frac{r^*}{n}\right)\left(\frac{1}{r}-\frac{1}{N}\right)\left(\left(\lambda\hat{\beta}\bar{X}+\lambda^2\bar{Y}+\frac{\hat{\beta}\bar{X}}{2}+\frac{\lambda\bar{Y}}{2}+\frac{3\bar{Y}}{8}\right)C_x^2-\left(\lambda\bar{Y}+\bar{Y}\right)\rho_{xy}C_xC_y\right) \tag{1.21}$$

To obtain the MSE of $t_2$, square (1.21) and take expectation of the result, the MSE of $t_2$ is obtained.

$$MSE\left(t_2\right)=\left(\frac{1}{r}-\frac{1}{N}\right)\left(\bar{Y}^2C_y^2+\left(1-\frac{r^*}{n}\right)\left(\hat{\beta}\bar{X}+\bar{Y}\left(\lambda+\frac{1}{2}\right)\right)\left[\left(\hat{\beta}\bar{X}+\bar{Y}\left(\lambda+\frac{1}{2}\right)\right)C_x^2-2\bar{Y}\rho_{xy}C_xC_y\right]\right) \tag{1.22}$$

**3.0 Empirical Study**

This section, presents simulation studies designed to evaluate the efficiency of the proposed estimators to assess the performance of existing estimators, a simulation was carried out. A population datasets of 1000 units was created using function from table 1.Then, Simple Random Sampling without Replacement (SRSWOR) was use to select 100 units 1000 times. The Biases, MSEs and PREs of the considered estimators were computed using (1.23), (1.24) and (1.25)

$$Bias(T)=\frac{1}{1000}\sum_{j=1}^{1000}\left(T-\bar{Y}\right) \tag{1.23}$$

$$MSEs(T)=\frac{1}{1000}\sum_{j=1}^{1000}\left(T-\bar{Y}\right) \tag{1.24}$$

$$PREs(T)=\frac{MSE(V_a)}{MSE(T)}\times 100 \tag{1.25}$$

where T are any of the proposed or existing estimators

**Table 1 Population used for Simulation Study**

| **Population** | **Auxiliary Variable (X)** | **Study Variable (Y)** |
|---|---|---|
| 1 | $X\sim beta(1,2)$ | $Y = 10 * X + e$ |
| 2 | $X\sim gamma(1,3)$ | |
| 3 | $X\sim \exp(1)$ | |
| 4 | $X\sim unif(1,10)$ | |

**Table 2: Biases, MSEs and PREs of Estimators and under population I (beta)**

| ESTIMATOR | BIAS | MSE | PRE | ESTIMATOR | BIAS | MSE | PRE |
|---|---|---|---|---|---|---|---|
| $\bar{y}_{gmlss}$ | -0.0694 | 0.0048 | 100 | $\bar{y}_{gmlss}$ | -0.0694 | 0.0048 | 100 |
| Proposed Estimator $t_1$ | | | | Proposed Estimator | | | |
| $t_{1(1)}$ | -0.04596 | 0.002112 | 228.44 | $t_{2(1)}$ | 0.3477572 | 0.00135189 | 357.0 |
| $t_{1(2)}$ | -0.05858 | 0.003432 | 140.60 | $t_{2(2)}$ | 0.327557 | 0.0024571 | 196.4 |
| $t_{1(3)}$ | -0.06181 | 0.003821 | 126.30 | $t_{2(3)}$ | 0.3223945 | 0.0027922 | 172.8 |
| $t_{1(4)}$ | -0.12867 | 0.016556 | 29.151 | $t_{2(4)}$ | 0.2154225 | 0.0145526 | 33.16 |
| $t_{1(5)}$ | -0.05816 | 0.003383 | 142.65 | $t_{2(5)}$ | 0.3282335 | 0.0024148 | 199.8 |
| $t_{1(6)}$ | -0.06235 | 0.00388 | 124.15 | $t_{2(6)}$ | 0.3215389 | 0.0028498 | 169.3 |
| $t_{1(7)}$ | -0.08781 | 0.00771 | 62.583 | $t_{2(7)}$ | 0.2807891 | 0.006274 | 76.92 |
| $t_{1(8)}$ | -0.05933 | 0.00352 | 137.07 | $t_{2(8)}$ | 0.3263568 | 0.0025331 | 190.5 |
| $t_{1(19)}$ | -0.05509 | 0.00303 | 158.99 | $t_{2(19)}$ | 0.3331449 | 0.0021186 | 227.81 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t_{1(20)}$ | -0.022929 | 0.00052 | 918.07 | $t_{2(20)}$ | 0.3846141 | 0.0001798 | 2683.7 |
| $t_{1(21)}$ | -0.054603 | 0.00298 | 161.87 | $t_{2(21)}$ | 0.3339343 | 0.0020728 | 232.84 |
| $t_{1(22)}$ | -0.067658 | 0.00457 | 105.43 | $t_{2(22)}$ | 0.3130449 | 0.0034535 | 139.75 |
| $t_{1(23)}$ | -0.065011 | 0.00422 | 114.19 | $t_{2(23)}$ | 0.3172813 | 0.0031451 | 153.45 |
| $t_{1(24)}$ | -0.06818 | 0.004649 | 103.81 | $t_{2(24)}$ | 0.3122036 | 0.0035168 | 137.25 |
| $t_{1(25)}$ | -0.06005 | 0.003606 | 133.8308 | $t_{2(25)}$ | 0.325214 | 0.0026065 | 185.16 |
| $t_{1(26)}$ | -0.06250 | 0.003906 | 123.53 | $t_{2(26)}$ | 0.321290 | 0.0028666 | 168.36 |
| $t_{1(27)}$ | -0.08276 | 0.006849 | 70.462 | $t_{2(27)}$ | 0.288877 | 0.0054882 | 87.941 |

**Table 3: Biases, MSEs and PREs of Estimators and under population II (Gamma)**

| ESTIMATOR | BIAS | MSE | PRE | ESTIMATOR | BIAS | MSE | PRE |
|---|---|---|---|---|---|---|---|
| $\bar{y}_{gmlss}$ | 0.10237 | 0.0104807 | 100 | $\bar{y}_{gmlss}$ | 0.10237 | 0.010480 | 100 |
| Proposed Estimator $t_1$ | | | | Proposed Estimator $t_2$ | | | |
| $t_{1(1)}$ | 0.065744 | 0.0043223 | 242.477 | $t_{2(1)}$ | 0.35330 | 0.0026917 | 389.360 |
| $t_{1(2)}$ | 0.083925 | 0.0070434 | 148.800 | $t_{2(2)}$ | 0.37947 | 0.0048550 | 215.873 |
| $t_{1(3)}$ | 0.089754 | 0.0080559 | 130.099 | $t_{2(3)}$ | 0.38786 | 0.0056827 | 184.430 |
| $t_{1(4)}$ | 0.092233 | 0.0085071 | 123.199 | $t_{2(4)}$ | 0.39143 | 0.0060545 | 173.106 |
| $t_{1(5)}$ | 0.089651 | 0.0080374 | 130.399 | $t_{2(5)}$ | 0.38771 | 0.0056675 | 184.925 |
| $t_{1(6)}$ | 0.091377 | 0.0083498 | 125.519 | $t_{2(6)}$ | 0.39019 | 0.0059247 | 176.897 |
| $t_{1(7)}$ | 0.093035 | 0.0086556 | 121.085 | $t_{2(7)}$ | 0.39258 | 0.0061772 | 169.667 |
| $t_{1(8)}$ | 0.091306 | 0.0083367 | 125.716 | $t_{2(8)}$ | 0.39009 | 0.0059139 | 177.219 |
| $t_{1(19)}$ | 0.079755 | 0.0063609 | 164.765 | $t_{2(19)}$ | 0.37347 | 0.004302 | 243.571 |
| $t_{1(20)}$ | 0.090708 | 0.0082281 | 127.377 | $t_{2(20)}$ | 0.38923 | 0.0058242 | 179.944 |
| $t_{1(21)}$ | 0.086815 | 0.0075369 | 139.057 | $t_{2(21)}$ | 0.38363 | 0.0052572 | 199.356 |
| $t_{1(22)}$ | 0.073968 | 0.005471 | 191.554 | $t_{2(22)}$ | 0.36514 | 0.0035919 | 291.787 |
| $t_{1(23)}$ | 0.081380 | 0.0066228 | 158.252 | $t_{2(23)}$ | 0.37581 | 0.0045141 | 232.176 |
| $t_{1(24)}$ | 0.081193 | 0.0065923 | 158.983 | $t_{2(24)}$ | 0.37554 | 0.0044895 | 233.449 |
| $t_{1(25)}$ | 0.0799445 | 0.0063911 | 163.988 | $t_{2(25)}$ | 0.373742 | 0.0043272 | 242.205 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t_{1(26)}$ | 0.0871056 | 0.0075873 | 138.133 | $t_{2(26)}$ | 0.384050 | 0.0052985 | 197.805 |
| $t_{1(27)}$ | 0.0907925 | 0.0082432 | 127.142 | $t_{2(27)}$ | 0.389357 | 0.0058369 | 179.5592 |

**Table 4: Biases, MSEs and PREs of Estimators and under population III (Exponential)**

| ESTIMATOR | BIAS | MSE | PRE | ESTIMATOR | BIAS | MSE | PRE |
|---|---|---|---|---|---|---|---|
| $\bar{y}_{gmlss}$ | 2.113235 | 4.465763 | 100 | $\bar{y}_{gmlss}$ | 2.113235 | 4.465763 | 100 |
| Proposed Estimator $t_1$ | | | | Proposed Estimator $t_2$ | | | |
| $t_{1(1)}$ | 1.60226 | 2.567238 | 173.952 | $t_{2(1)}$ | 0.370218 | 2.058313 | 216.962 |
| $t_{1(2)}$ | 1.802479 | 3.248931 | 137.453 | $t_{2(2)}$ | 0.443522 | 2.616755 | 170.660 |
| $t_{1(3)}$ | 1.800959 | 3.243453 | 137.685 | $t_{2(3)}$ | 0.442966 | 2.612263 | 170.953 |
| $t_{1(4)}$ | 1.917439 | 3.676574 | 121.465 | $t_{2(4)}$ | 0.485611 | 2.967652 | 150.481 |
| $t_{1(5)}$ | 1.839543 | 3.383917 | 131.970 | $t_{2(5)}$ | 0.457092 | 2.727475 | 163.732 |
| $t_{1(6)}$ | 1.774492 | 3.14882 | 141.823 | $t_{2(6)}$ | 0.433275 | 2.534669 | 176.187 |
| $t_{1(7)}$ | 1.900243 | 3.610924 | 123.673 | $t_{2(7)}$ | 0.479315 | 2.913759 | 153.264 |
| $t_{1(8)}$ | 1.813763 | 3.289738 | 135.748 | $t_{2(8)}$ | 0.447654 | 2.650222 | 168.505 |
| $t_{1(19)}$ | 1.777506 | 3.159528 | 141.342 | $t_{2(19)}$ | 0.434379 | 2.543448 | 175.579 |
| $t_{1(20)}$ | 1.901305 | 3.61496 | 123.535 | $t_{2(20)}$ | 0.479704 | 2.917073 | 153.09 |
| $t_{1(21)}$ | 1.81527 | 3.295203 | 135.523 | $t_{2(21)}$ | 0.448205 | 2.654705 | 168.220 |
| $t_{1(22)}$ | 1.665817 | 2.774946 | 160.931 | $t_{2(22)}$ | 0.393487 | 2.22833 | 200.408 |
| $t_{1(23)}$ | 1.664983 | 2.772167 | 161.092 | $t_{2(23)}$ | 0.393182 | 2.226055 | 200.613 |
| $t_{1(24)}$ | 1.689513 | 2.854456 | 156.448 | $t_{2(24)}$ | 0.402163 | 2.293446 | 194.718 |
| $t_{1(25)}$ | 1.739643 | 3.026356 | 147.562 | $t_{2(25)}$ | 0.420516 | 2.434286 | 183.452 |
| $t_{1(26)}$ | 1.738249 | 3.021509 | 147.799 | $t_{2(26)}$ | 0.420006 | 2.430314 | 183.752 |
| $t_{1(27)}$ | 1.870916 | 3.500326 | 127.581 | $t_{2(27)}$ | 0.468578 | 2.822988 | 158.192 |

**Table 5: Biases, MSEs and PREs of Estimators and under population IV (Uniform)**

| ESTIMATOR | BIAS | MSE | PRE | ESTIMATOR | BIAS | MSE | PRE |
|---|---|---|---|---|---|---|---|
| $\bar{y}_{gmlss}$ | 1.040 | 0.680506 | 100 | $\bar{y}_{gmlss}$ | 1.0405 | 0.680506 | 100 |
| Proposed Estimator $t_1$ | | | | Proposed Estimator $t_2$ | | | |
| $t_{1(1)}$ | 0.6838 | 0.137023 | 496.6369 | $t_{2(1)}$ | -0.847047 | 0.02948791 | 2307.749 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t_{1(2)}$ | 0.8676 | 0.241331 | 281.9803 | $t_{2(2)}$ | 0.762456 | 0.09015195 | 754.8443 |
| $t_{1(3)}$ | 0.8721 | 0.245890 | 276.7516 | $t_{2(3)}$ | -0.760385 | 0.09312021 | 730.7832 |
| $t_{1(4)}$ | 0.9257 | 0.4643314 | 146.5563 | $t_{2(4)}$ | -0.735732 | 0.2521119 | 269.9226 |
| $t_{1(5)}$ | 0.8756 | 0.383333 | 177.5237 | $t_{2(5)}$ | -0.75877 | 0.1900951 | 357.9823 |
| $t_{1(6)}$ | 0.8339 | 0.3282938 | 207.286 | $t_{2(6)}$ | -0.77798 | 0.149762 | 454.3922 |
| $t_{1(7)}$ | 0.8984 | 0.5037856 | 135.0787 | $t_{2(7)}$ | -0.748316 | 0.2832483 | 240.251 |
| $t_{1(8)}$ | 0.8377 | 0.4555375 | 149.3855 | $t_{2(8)}$ | -0.776229 | 0.2452482 | 277.4768 |
| $t_{1(19)}$ | 0.8242 | 0.3166889 | 214.8818 | $t_{2(19)}$ | -0.78242 | 0.1414849 | 480.975 |
| $t_{1(20)}$ | 0.8945 | 0.501948 | 135.5732 | $t_{2(20)}$ | -0.750109 | ] 0.281786 | 241.4975 |
| $t_{1(21)}$ | 0.8329 | 0.4517067 | 150.6524 | $t_{2(21)}$ | -0.778446 | 0.2422674 | 280.8909 |
| $t_{1(22)}$ | 0.7608 | 0.1625728 | 418.5861 | $t_{2(22)}$ | -0.811589 | 0.0427614 | 1591.405 |
| $t_{1(23)}$ | 0.7646 | 0.1639806 | 414.9922 | $t_{2(23)}$ | -0.809851 | 0.0435296 | 1563.319 |
| $t_{1(24)}$ | 0.7677 | 0.2307597 | 294.8985 | $t_{2(24)}$ | -0.808440 | 0.0833499 | 816.4453 |
| $t_{1(25)}$ | 0.8204 | 0.2010124 | 338.5398 | $t_{2(25)}$ | -0.784168 | 0.0648742 | 1048.962 |
| $t_{1(26)}$ | 0.8252 | 0.2041937 | 333.2654 | $t_{2(26)}$ | -0.78195 | 0.0667992 | 1018.735 |
| $t_{1(27)}$ | 0.8913 | 0.4205507 | 161.8133 | $t_{2(27)}$ | 0.7515617 | 0.218242 | 311.813 |

Table 2-5present numerical results that compare the biases, mean squared errors (MSEs) and percentage relative efficiencies (PREs) of existing estimators and the new estimators proposed in the study. The findings indicate that all the proposed estimators exhibit lower MSEs and higher PREs compared to the existing estimators considered in the investigation.The results showed that the proposed estimators $t_1$ and $t_2$ exhibited lower MSEs with significant percentage gains in efficiency compared to the estimator proposed by Subramani and Gupta (2014) in all the cases considered for the empirical studies with the exception of few cases where few members of the new estimators of the proposed performed below standard.

## 4.0 Conclusion

The empirical investigation revealed that the proposed estimators consistently are efficient than the estimator introduced by Subramani and Gupta (2014) in terms of efficiency, across all the scenarios examined. Notably, the proposed estimators demonstrated superior performance in the vast majority of cases, with only few exceptions where a subset of the new estimators exhibited subpar results. These findings highlight the proposed methodology’s robustness and accuracy advantages, making it valuable contribution to estimator development.

**REFERENCES**


Audu, A., Singh,R., & Supriya, K.(2021). New regression-type compromised imputation class of estimators with known parameters of auxiliary variable. *Journal of Statistics and Management Systems*,52(10), 4789-4801. DOI:10.1080/03610918.2021.197082

Audu, A. & Singh, R. V. K. (2021). Exponential-type regression compromised imputation class of estimators . *Journal of Statistics and Management Systems*, 24(6), 1253-1266.

Ayed, A. E. & Khalid, U. I. (2023). A New exponential type ratio estimator for the population mean in systematic sampling, 18 (74), 422-449.

Azeem, M & Khan, Z. (2021). An Improved Diagonal-Cum-Linear Systematic Sampling scheme. *Indian Journal of Economics and Business*, 20(2), 1197-1209.

Azeem, M., Hussain, S., Ijaz, M., & Salahuddin, N. (2023). An improved version of systematic sampling design for use with linear trend data .DOI:10.16/j.heliyon.2023e17121.

Chang, H. J. & Hauang, K.C. (2000). Remainder linear systematic sampling, *Sankhya:Indian Journal Of Statistic*, 62, 249-256.

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

Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ration to produce. *Journal of Agricultural Science*, 30, 262-275.

Gautschi, W. (1957). Some remarks on systematic sampling, *Annals of Mathematical Statistics*. 28, 385–394.

Hajeck, J. (1959). Optimum strategy and other problems in probability sampling, *Casopis propestovani Matematikly*, 84, 387-423.

Hansen, M. H. Hurwitz, W. N. Gurney, M. (1946). Problems and methods of the sample survey of business, *The Journal of theAmerican Statistical Association*, 41(234), 173–189.

Kadilar, C. & Cingi, H. (2006). Improvement in estimating the population mean in simple random sampling. *Applied Mathematics Letters*, 19(1), 75-79.

Kadilar, C. and Cingi, H. (2008). Estimators for the population mean in the case of missing data, Communications in Statistics- Theory and Methods, 37, 2226-2236.

Khan, M. S., Ullah, A., Al-Hossain. Y., & Bashir, N. (2015). Improved ratio-type estimators using maximum and minimum values under simple random sampling scheme, Hacette. *Journal of Mathematics and Statistics*, 44(4), 923–931.

Khan, M. & Rajesh, S. (2015). Estimation of population mean in chain ratio-type estimator under systematic sampling, *journal of probability and statistics.* DOI:10.1155/2015/248374.

Kushwaha, K. S. & Singh, H. P. (1989). Class of almost unbiased ratio and product estimators in systematic sampling, *Journal of the Indian Society of AgriculturalStatistics.* 2(41), 193-205.

Lahiri, D. B. (1951). A method for selection providing unbiased estimates. *International Statistical Association Bulletin*, 33, 133-140.

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

Noor-ul-Amin, M., Javiad, A., Hanif, M. (2017). Estimation of population mean in systematic Pestovani Matematiky, 84, 387– 423.

Robson, D. S. (1957). Application of multivariable polykas to the theory of un -biased ratio type estimation. *Journal of American Statistics Associations,* 50, 1225-1226.

Singh, S. and Deo, B. (2003). Imputation by power transformation. Statistical Papers, 44, 555-579. Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. Metrika, 51, 267-s276

Sampath, S. & Varalakshmi, V. (2008). Diagonal circular systematic sampling, Model AssistedStatistics Application. 3, 345–352.

Shukla, N. D., (1971). Systematic sampling and product method of estimation. *In Processing of All India Seminar on Demography and Statistics*.BHU,Varansi, India.

Signh, H. P., & Tailor, R. (2003). Use of known correlation coefficient in estimating the finite. the population mean. *Statistics in Transition,* 6, 555-560.

Singh, H. P. & Solanki, R. S. (2012). An efficient class of estimators for the population mean using auxiliary information in systematic sampling. *Journal of Statistical Theory and Practice,* 6(2), 274–285.

Singh, H. P. Tailor, R. Jatwa, N. K. (2011). Modified ratio and product estimators for population mean in systematic sampling, *Journal of Modern Applied Statistical Methods*, 10(2),424–435.

Singh, H. P., & Jatwa, N.K (2012). A class of exponential-type estimators in systematic sampling. *Economic Quality Control*, 27, 915-208.

Singh, M. P. (1967). Ratio cum product method of estimation, Metrika, 12(1), 34–42.

Singh, R. Malik, S. Chaudhary, M. K. Verma, H. & 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, *Questiio*, 22(3), 403–416.

Singh, R. Chauhan. P., Sawan. N.,& Smarandache, F. (2011).Improved exponential estimatorfor population variance using two auxiliary variables,*Italian Journal of Pure andApplied Mathematics*. 28, 101–108.

Singh, R., Malik., S. Singh., V. K. (2012). An improved estimator in systematic sampling. *Journal Of Scientific Research*, 56, 177-182.

Subramani, J. (2000). Diagonal systematic sampling scheme for finite populations,*Journal of theIndian Society of Agricultural Statistics*, 53(2), 187-195.

Subramani, J. (2009). Further results on diagonal systematic sampling for finite populations, *Journal of Indian Society of Agricultural Statistic*, 63, 277-282.

Subramani, J., Gupta, S.N.(2014). Generalized modified linear systematic sampling scheme forfinite populations, *Hacet. Journal of Mathematics and Statistic*, 43, 529-542.

Swain, A. K. P. C. (1964). The use of systematic sampling ratio estimate. *Journal of IndianStatistical Association*, 2, 160-164.

Toutenburg, H., Srivastava, V. K. and Shalabh, A. (2008). Imputation versus imputation of missing values through ratio method in sample surveys. *Statistical Papers*, 49, 237-247.

Tailor, T. Jatwa., N. K. Singh., H. P.( 2013). A ratio-cum-product estimator of finite population mean in systematic sampling, *Statistics In Transition.* 14(3), 391-398.

Verma, H. K. & Singh, R. (2014). A family of efficient estimator in circular systematic sampling,*Journal of Advanced Computing*,3(2), 56–68.

Verma, H. K. Singh, R. D. Singh, R. (2014), Some improved estimators in systematic sampling. under non-response, *National Academy Science Letters*, 37(1), 91–95.

Wang, L. and Wang, Q., (2006). Empirical likelihood for parametric model under imputation for missing data. *Journal of Statistics and Management Systems*, 9 (1), 1-13.

Zahoor, A. G., Tanveer A. Y.,&Khalid, I. (2022). A new data gathering exponential type-ratio estimator for the Population Mean.*Journal of Scientific Research & Reports*28(10), 148-153. DOI: 10.9734202228101697.