Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(6): 9825–9834.
DOI: 10.3934/math.2022547
Received: 17 January 2022
Revised: 26 February 2022
Accepted: 06 March 2022
Published: 18 March 2022

Research article

Estimation of finite population mean in presence of maximum and minimum values under systematic sampling scheme

Anum Iftikhar[1], Hongbo Shi[1], Saddam Hussain[2,*], Ather Qayyum[2], M. El-Morshedy[3,4] and Sanaa Al-Marzouki[5]

[1] School of Statistics, Shanxi University of Finance and Economics, Taiyuan 030031, China
[2] Department of Mathematics and Statistics, Institute of Southern Punjab Multan, Pakistan
[3] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[4] Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
[5] Statistics Department, Faculty of Science, King AbdulAziz University, Jeddah 21551, Saudi Arabia

* **Correspondence:** Email: saddamhussain.stat885@gmail.com.

Abstract: Estimators for the finite population mean of the research variable are proposed in this article, employing ratio, product, and regression type estimators, all of which need just one auxiliary variable. A first-order approximation is developed for the mean squared errors of the techniques provided. It has been proven theoretically that the suggested estimators perform better than current estimators, and these theoretical conditions have been validated numerically using four data sets.

Keywords: systematic sampling; bias; efficiency; mean square error (MSE); maximum and minimum values
Mathematics Subject Classification: 03F20, 00A71

1. Introduction

Auxiliary information may enhance the accuracy of estimators in survey sampling, as is well known. It is possible to employ the supplementary information in the selection and estimate stages. Many academics have sought to acquire population metrics like the mean or median that have the best statistical qualities. For this reason, a representative sample of the population is required. Simple random sampling (SRS) may pick the units when the population of interest is homogenous. However, it is not easy to use the SRS or any other sampling strategy to estimate population characteristics with a natural population, such as a forest. Systematic sampling may be used swiftly in this case to choose a representative sample from the population.

Systematic sampling benefits picking the whole sample with a single random start, making it extremely simple to operate. Apart from its simplicity, which is critical in large-scale sampling operations, it produces more efficient estimators than those produced by SRS or stratified random sampling for specific populations. References [11, 14] proposed ratio and product estimators for estimating the finite population mean $\bar{Y}$ of the research variable y, respectively, along with their systematic sampling features. By contrast, systematic sampling has been extensively addressed by [1, 2, 4–8, 13].

2. Materials and methods

It is important to remember that a finite population with units $U = [1...N]$ has the study variable y and the auxiliary variable x. Every kth unit is randomly picked from a systematic sample of size n, with a random start to choose the initial unit. $N = nk$, where n and k are positive integers, are used in this example. There are y_{ij} and x_{ij}, which are the values of the j^{th} unit in the i^{th} chosen sample for y and x variables, respectively, in the i^{th}-selected sample. Systematic random sampling uses sample means such as these:

$$\bar{y}_{sys} = \frac{1}{n}\sum_{j=1}^{n} y_{ij},\ \ \bar{x}_{sys} = \frac{1}{n}\sum_{j=1}^{n} x_{ij},$$

which are unbiased estimators for population means $\bar{Y}$ and $\bar{X}$ respectively. To obtain the biases and mean square errors, we define:

$$e_0 = \frac{\bar{y}_{sys} - \bar{Y}}{\bar{Y}},\ \ e_1 = \frac{\bar{x}_{sys} - \bar{X}}{\bar{X}}, \lambda = \frac{1}{n} - \frac{1}{N}$$

such that, $E(e_0) = E(e_1) = 0$

$$E(e_0^2) = \lambda\rho_y^{*2}C_y^2,\ \ E(e_1^2) = \lambda\rho_x^{*2}C_x^2,\ \ E(e_0e_1) = \lambda\rho_{yx}C_yC_x\sqrt{\rho_y^*\rho_x^*},$$

where

$$\rho_y^* = 1 + (n-1)\rho_y,\ \ \rho_x^* = 1 + (n-1)\rho_x,$$

$$C_y = \frac{S_y}{\bar{Y}},\ \ \ S_y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y}^2)},$$

$$C_x = \frac{S_x}{\bar{X}},\ \ \ S_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{X}^2)},$$

C_y, C_x are the coefficients of Y and X respectively, and

$$\rho_y = \frac{E[(y_{ij} - \bar{Y})(\acute{y}_{ij} - \bar{Y})]}{E(y_{ij} - \bar{Y})} \text{ and } \rho_x = \frac{E[(x_{ij} - \bar{X})(\acute{x}_{ij} - \bar{X})]}{E(x_{ij} - \bar{X})},$$

correlation coefficients within a single systematic sample between two or more individuals,

$$\rho_{yx} = \frac{E(y_{ij} - \bar{Y})(x_{ij} - \bar{X})}{\sqrt{E(y_{ij} - \bar{Y})^2 E(x_{ij} - \bar{X})^2}},$$

is the correlation between y and x.

Reference [10] suggested the following an unbiased estimator:

$$\bar{y}_s = \begin{cases} \bar{y} + c, & \text{if sample contains } y_{min} \text{ but not } y_{max}, \\ \bar{y} - c, & \text{if sample contains } y_{max} \text{ but not } y_{min}, \\ \bar{y}, & \text{for all samples.} \end{cases} \quad (2.1)$$

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

$$V(\bar{y}_s) = \lambda S_y^2 - \frac{2\lambda nc}{N-1}(y_{max} - y_{min} - nc), \quad (2.2)$$

where S_y^2 is the population variance, and c is the constant. The optimum value of c is,

$$c_{opt} = \frac{(y_{max} - y_{min})^2}{2n}.$$

The minimum mean square error of $\bar{y}_s$:

$$V(\bar{y}_s)_{min} = V(\bar{y}) - \frac{\lambda(y_{max} - y_{min})^2}{2(N-1)}. \quad (2.3)$$

The minimum variance of $\bar{y}_s$ is always smaller than the variance of $\bar{y}$.

The usual ratio and product estimators in systematic sampling scheme are given by:

$$\bar{y}_{Rsys} = \bar{y}_{sys}\left(\frac{\bar{X}}{\bar{x}_{sys}}\right) \quad (2.4)$$

$$\bar{y}_{Psys} = \bar{y}_{sys}\left(\frac{\bar{x}_{sys}}{\bar{X}}\right). \quad (2.5)$$

Biases and first-degree approximation mean square errors of $\bar{y}_{Rsys}$ and $\bar{y}_{Psys}$ are, respectively, provided by:

$$B(\bar{y}_{Rsys}) = \lambda\bar{Y}(C_x^2\rho_x^* - \rho_{yx}C_yC_x\sqrt{\rho_y^*\rho_x^*}), \quad (2.6)$$

$$B(\bar{y}_{Psys}) = \lambda\bar{Y}(\rho_{yx}C_yC_x\sqrt{\rho_y^*\rho_x^*}), \quad (2.7)$$

and

$$MSE(\bar{y}_{Rsys}) = \lambda\left[S_y^*\rho_y^* + R^2S_x^*\rho_x^* - 2RS_{yx}\sqrt{\rho_y^*\rho_x^*}\right], \quad (2.8)$$

$$MSE(\bar{y}_{Psys}) = \lambda\left[S_y^*\rho_y^* + R^2S_x^*\rho_x^* + 2RS_{yx}\sqrt{\rho_y^*\rho_x^*}\right]. \quad (2.9)$$

The usual regression estimator under systematic random sampling scheme, is given by:

$$\bar{y}_{lrsys} = \bar{y}_{sys} + b(\bar{X} - \bar{x}_{sys}). \quad (2.10)$$

The sample regression coefficient b is represented by this equation. b is the least square estimates of the population regression coefficient (β), therefore $\bar{y}_{lrsys}$, has the variance provided by:

$$V(\bar{y}_{lrsys}) = \lambda\rho_y^*S_y^2(1 - \rho_{yx}^2). \quad (2.11)$$

3. Purposed estimator

On the lines of [10], we provide estimators of the ratio, product, and regression types in the presence of maximum and minimum values using a systematic sampling approach. As follows, we propose a better estimator for each of the following situations.

3.1. Case (1): When correlation between y and x is positive

When y and x have a positive correlation, it is considered that a higher value of x equals a much higher value of y. A lower value of x is regarded to be equivalent to a smaller value of y when y equals x. As a consequence, the following ratio type estimator is defined.

$$\hat{\hat{Y}}_{R(sys)} = \bar{y}_{sys(c_{11})}\frac{\bar{X}}{\bar{x}_{sys(c_{21})}}$$

or

$$\hat{\hat{Y}}_{R(sys)} = \begin{cases} \frac{\bar{y}_{sys}+c_1}{\bar{x}_{sys}+c_2}\bar{X} \\ \frac{\bar{y}_{sys}-c_1}{\bar{x}_{sys}-c_2}\bar{X} \\ \frac{\bar{y}_{sys}}{\bar{x}_{sys}}\bar{X} \end{cases}. \tag{3.1}$$

The regression type estimator is:

$$\hat{\hat{Y}}_{lr1(sys)} = \bar{y}_{sys(c_{11})} + b(\bar{X} - \bar{x}_{sys(c_{21})}), \tag{3.2}$$

where b is the sample regression coefficient.

Consider $(\bar{y}_{sys(c_{11})} = \bar{y}_{sys} + c_1, \bar{x}_{sys(c_{21})} = \bar{x}_{sys} + c_2)$ if the sample contains y_{min} and x_{min}, $(\bar{y}_{sys(c_{11})} = \bar{y}_{sys} - c_1, \bar{x}_{sys(c_{21})} = \bar{x}_{sys} - c_2)$ if the sample contains y_{max} and x_{max} and $(\bar{y}_{sys(c_{11})} = \bar{y}_{sys}, \bar{x}_{sys(c_{21})} = \bar{x}_{sys})$ for all the other combinations of sample.

3.2. Case (2): When correlation between y and x is negative

When y and x have a negative correlation, it is considered that a higher value of x equals a much lower value of y. A lower value of x is regarded to be equivalent to a smaller value of y when y equals x. As a consequence, the following ratio type estimator is defined.

$$\hat{\hat{Y}}_{P(sys)} = \bar{y}_{sys(c_{12})}\frac{\bar{x}_{sys(c_{22})}}{\bar{X}} \tag{3.3}$$

or

$$\hat{\hat{Y}}_{P(sys)} = \begin{cases} \frac{(\bar{y}_{sys}+c_1)(\bar{x}_{sys}-c_2)}{\bar{X}} \\ \frac{(\bar{y}_{sys}-c_1)(\bar{x}_{sys}+c_2)}{\bar{X}} \\ \frac{\bar{y}_{sys}}{\bar{x}_{sys}}\bar{X} \end{cases}. \tag{3.4}$$

The regression type estimator is:

$$\hat{\hat{Y}}_{lr2(sys)} = \bar{y}_{sys(c_{11})} + b(\bar{X} - \bar{x}_{sys(c_{22})}), \tag{3.5}$$

where b is the sample regression coefficient.

Consider ($\bar{y}_{sys(c_{12})} = \bar{y}_{sys} + c_1, \bar{x}_{sys(c_{22})} = \bar{x}_{sys} - c_2$) if the sample contains y_{min} and x_{min}, ($\bar{y}_{sys(c_{12})} = \bar{y}_{sys} - c_1, \bar{x}_{sys(c_{22})} = \bar{x}_{sys} + c_2$) if the sample contains y_{max} and x_{max} and ($\bar{y}_{sys(c_{12})} = \bar{y}_{sys}, \bar{x}_{sys(c_{22})} = \bar{x}_{sys}$) for all the other combinations of sample. Let

$$e_0 = \frac{\bar{y}_{sys(c_1)} - \bar{Y}}{\bar{Y}}, \qquad e_1 = \frac{\bar{x}_{sys(c_1)} - \bar{X}}{\bar{X}},$$

$$E(e_0^2) = \left(\frac{\lambda}{\bar{Y}^2}\right)\left[S_y^2\rho_y^* - \frac{2nc_1}{N-1}\left(\Delta_y - nc_1\right)\right],$$

$$E(e_1^2) = \left(\frac{\lambda}{\bar{X}^2}\right)\left[S_x^2\rho_x^* - \frac{2nc_2}{N-1}\left(\Delta_x - nc_2\right)\right],$$

$$E(e_0e_1) = \left(\frac{\lambda}{\bar{Y}\bar{X}}\right)\left[S_{uv}\sqrt{\rho_y^*\rho_x^*} - \frac{n}{N-1}\left\{c_2\Delta_y + c_1\Delta_x - 2nc_1c_2\right\}\right],$$

where

$$\Delta_y = (y_{max} - y_{min}),\ \ \Delta_x = (x_{max} - x_{min}).$$

Expressing $\hat{\hat{Y}}_{R(sys)}$ in terms of e's,

$$\hat{\hat{Y}}_{R(sys)} = \bar{Y}(1 + e_0)(1 + e_1)^{-1}. \tag{3.6}$$

By right hand side of (3.6) up-to first order of approximation, we have

$$(\hat{\hat{Y}}_{R(sys)} - \bar{Y}) = \bar{Y}(e_0 + e_1 - e_0e_1 + e_1^2). \tag{3.7}$$

Using the above equation, the bias of $\hat{\hat{Y}}_{R(sys)}$, is given by:

$$\begin{aligned} B(\hat{\hat{Y}}_{R(sys)}) \cong & \frac{\lambda}{\bar{X}}\Big[(RS_x^2\rho_x^* - S_{yx}\sqrt{\rho_y^*\rho_x^*}) - \frac{n}{N-1}\Big\{2c_2(\Delta_x - nc_2) \\ & +(c_2\Delta_y + c_1\Delta_x - 2nc_1c_2)\Big\}\Big]. \end{aligned} \tag{3.8}$$

Using (3.7), the mean error of $\hat{\hat{Y}}_{R(sys)}$ up to the first order of approximation, is given by:

$$\begin{aligned} MSE(\hat{\hat{Y}}_{R(sys)}) \cong & \lambda\Big[S_y^2\rho_y^* - \frac{2nc_1}{N-1}\left(\Delta_y - nc_1\right) + R^2\left\{S_x^2\rho_x^* - \frac{2nc_2}{N-1}\left(\Delta_x - nc_2\right)\right\} \\ & - 2R\left\{S_{yx}\sqrt{\rho_y^*\rho_x^*} - \frac{n}{N-1}\left(c_2\Delta_y + c_1\Delta_x - 2nc_1c_2\right)\right\}\Big], \end{aligned} \tag{3.9}$$

where $R = \frac{\bar{Y}}{\bar{X}}$, or

$$\begin{aligned} MSE(\hat{\hat{Y}}_{R(sys)}) \cong & \Big[S_y^2\rho_y^* + R^2S_x^2\rho_x^* - 2RS_{yx}\sqrt{\rho_y^*\rho_x^*}\Big] - \frac{2n\lambda}{N-1}\Big\{c_1(\Delta_y - nc_1) \\ & +c_2R^2(\Delta_x - nc_2) - R(c_2\Delta_y + c_1\Delta_x - 2nc_1c_2)\Big\} \end{aligned} \tag{3.10}$$

or

$$MSE(\hat{\hat{Y}}_{R(sys)}) \cong MSE(\bar{y}_{R_{sys}}) - \frac{2\lambda n}{N-1}\Big[(c_1 - Rc_2)\left\{\Delta_y - R\Delta_x - n\left(c_1 - Rc_2\right)\right\}\Big]. \tag{3.11}$$

Optimum values of c_1 and c_2, we differentiate (3.11) with respect to c_1 and c_2 as:

$$\frac{\partial M(\hat{\hat{Y}}_{R(sys)})}{\partial c_1} = \Delta_y - R\Delta_x - 2n(c_1 - Rc_2) = 0,$$
$$\frac{\partial M(\hat{\hat{Y}}_{R(sys)})}{\partial c_2} = \Delta_y - R\Delta_x - 2n(c_1 - Rc_2) = 0.$$

As a result, a one-of-a-kind solution is not attainable. The ideal values of c_1 and c_2 are:

$$c_{1opt} = \frac{\Delta_y}{2n} \text{ and } c_{2opt} = \frac{\Delta_x}{2n}.$$

For optimum values of c_1 and c_2, the minimum MSE of $\hat{\hat{Y}}_{R(sys)}$, is given as:

$$MSE(\hat{\hat{Y}}_{R(sys)})_{min} \cong MSE(\bar{y}_{R(sys)}) - \frac{\lambda}{2(N-1)}\left[\Delta_y - R\Delta_x\right]^2. \tag{3.12}$$

Similarly the bias and mean squared error of $\hat{\hat{Y}}_{P(sys)}$ are respectively, given by:

$$B(\hat{\hat{Y}}_{P(sys)}) \cong \frac{\lambda}{\bar{X}}\left[S_{yx}\sqrt{\rho_y^*\rho_x^*} - \frac{n}{N-1}\left\{\left(c_2\Delta_y + c_1\Delta_x - 2nc_1c_2\right)\right\}\right] \tag{3.13}$$

and

$$MSE(\hat{\hat{Y}}_{P(sys)}) \cong MSE(\bar{y}_{P(sys)}) - \frac{2\lambda n}{N-1}\Big[(c_1 + Rc_2)\{\Delta_y + R\Delta_x - n(c_1 + Rc_2)\}\Big]. \tag{3.14}$$

For optimum values of c_1 and c_2, the MSE of $\hat{\hat{Y}}_{P(sys)}$ is given as:

$$MSE(\hat{\hat{Y}}_{P(sys)})_{min} \cong MSE(\bar{y}_{P(sys)}) - \frac{\lambda}{2(N-1)}\left[\Delta_y - R\Delta_x\right]^2. \tag{3.15}$$

The variance of regression estimator $\hat{\hat{Y}}_{lr1(sys)}$ in case of positive correlation, is given by:

$$V(\hat{\hat{Y}}_{lr1(sys)}) = V(\bar{y}_{lr(sys)}) - \frac{2\lambda n}{N-1}\Big[(c_1 - \beta c_2)\Big\{\Delta_y - \beta\Delta_x - 2n(c_1 - \beta c_2)\Big\}\Big], \tag{3.16}$$

where $\beta = \rho_{yx}\frac{S_y\rho_y^*}{S_x\rho_x^*}$ is the population regression coefficient of y on x . For

$$c_{1opt} = \frac{\Delta_y}{2n} \text{ and } c_{2opt} = \frac{\Delta_x}{2n},$$

the minimum variance of ($\hat{\hat{Y}}_{lr1sys}$), is given as:

$$V(\hat{\hat{Y}}_{lr1(sys)}) = V(\bar{y}_{lr(sys)}) - \frac{\lambda}{2(N-1)}\left[\Delta_y - \beta\Delta_x\right]^2. \tag{3.17}$$

For negative correlation, variance of the regression estimator ($\hat{\hat{Y}}_{lr2(sys)}$), is given by:

$$V(\hat{\hat{Y}}_{lr2(sys)}) = V(\bar{y}_{lr(sys)}) - \frac{2\lambda n}{N-1}\Big[(c_1 + \beta c_2)\Big\{\Delta_y + \beta\Delta_x - 2n(c_1 + \beta c_2)\Big\}\Big]. \tag{3.18}$$

For optimum values of c_1 and c_2 *i.e*:

$$c_{1opt} = \frac{\Delta_y}{2n} \text{ and } c_{2opt} = \frac{\Delta_x}{2n}.$$

The minimum variance of $(\bar{Y}_{lr2(sys)})$, is given by:

$$V(\hat{\bar{Y}}_{lr2(sys)})_{min} = V(\bar{y}_{lr(sys)}) - \frac{\lambda}{2(N-1)}\left[\Delta_y + \beta\Delta_x\right]^2. \tag{3.19}$$

In general we can write

$$V(\hat{\bar{Y}}_{lr(g)(sys)})_{min} = V(\bar{y}_{lr(sys)}) - \frac{\lambda}{2(N-1)}\left[\Delta_y - |\beta|\,\Delta_x\right]^2. \tag{3.20}$$

4. Comparison of estimators

This section compares the proposed estimators to the conventional ratio, product, and regression type estimators in systematic sampling.

4.1. Condition (i)

The purposed ratio type estimator $\hat{\bar{Y}}_{R(sys)}$ systematic sampling will outperform the conventional ratio type estimator, by (2.8) and (3.12), if

$$\left[MSE(\bar{y}_{R(sys)}) - MSE(\hat{\bar{Y}}_{R(sys)})_{min}\right] \geq 0$$

or

$$min\left[Rc_2, Rc_2 - \left\{\frac{R\Delta_x - \Delta_y}{n}\right\}\right] < c_1 < max\left[Rc_2, Rc_2 - \left\{\frac{R\Delta_x - \Delta_y}{n}\right\}\right]. \tag{4.1}$$

4.2. Condition (ii)

The purposed product type estimator $\hat{\bar{Y}}_{P(sys)}$ systematic sampling will outperform the conventional ratio type estimator, by (2.9) and (3.15), if

$$\left[MSE(\bar{y}_{P(sys)}) - MSE(\hat{\bar{Y}}_{P(sys)})_{min}\right] \geq 0$$

or

$$min\left[-Rc_2, -Rc_2 + \left\{\frac{R\Delta_x + \Delta_y}{n}\right\}\right] < c_1 < max\left[-Rc_2, -Rc_2 + \left\{\frac{R\Delta_x + \Delta_y}{n}\right\}\right]. \tag{4.2}$$

4.3. Condition (iii)

The proposed regression type estimator $\hat{\bar{Y}}_{lr1(sys)}$ better than the standard regression estimator in the case of random sampling, by (2.11) and (3.17), if

$$\left[V(\bar{y}_{lr(sys)}) - V(\hat{\bar{Y}}_{lr1(sys)})\right] \geq 0$$

or

$$min\left[\beta c_2, \beta c_2 - \left\{\frac{\beta\Delta_x - \Delta_y}{n}\right\}\right] < c_1 < max\left[\beta c_2, \beta c_2 - \left\{\frac{\beta\Delta_x - \Delta_y}{n}\right\}\right]. \tag{4.3}$$

4.4. Condition (iv)

The proposed regression type estimator $\hat{\hat{Y}}_{lr2(sys)}$ will perform better than the usual regression type estimator in systematic sampling, by (2.11) and (3.19).

$$\left[V(\bar{y}_{lr(sys)}) - V(\hat{\hat{Y}}_{lr2(sys)})\right] \geqslant 0$$

or

$$min\left[-\beta c_2, -\beta c_2 + \left\{\frac{\beta\Delta_x - \Delta_y}{n}\right\}\right] < c_1 < max\left[-\beta c_2, -\beta c_2 + \left\{\frac{\beta\Delta_x - \Delta_y}{n}\right\}\right]. \quad (4.4)$$

If the conditions outlined in $(i) - (iv)$ are met, the suggested estimators outperform the current ones.

5. Numerical study

In this part, we compare the recommended estimators against various other estimators using four different sets of data sets. The following are the population data descriptions that are required.

Population 1: [Source: [12]].

First 32 observations are considered as a population.

y = Cultivated area in acres in 1974 census,

x =Cultivated area in 1971 census.

Data statistics are given as:

$N = 32$, $n = 10$, $\bar{Y} = 199.2813$, $\bar{X} = 207.8438$, $C_y = 0.7738$, $C_x = 0.7459$, $\rho_y^* = 0.0235$, $\rho_x^* = 0.0144$, $\rho_{yx} = 0.9815$, $y_{max} = 634$, $y_{min} = 6$, $x_{max} = 564$, $x_{min} = 5$.

Population 2: [Source: PSLM(2007-2008)].

y= Age of students.

x=Total monthly expenditure on education.

Data statistics are given as:

$N = 144$, $n = 70$, $\bar{Y} = 18.0833$, $\bar{X} = 4200.1390$, $C_y = 0.6738$, $C_x = 0.8459$, $\rho_y^* = 0.0235$, $\rho_x^* = 0.0144$, $\rho_{yx} = 0.7815$, $y_{max} = 80$, $y_{min} = 1$, $x_{max} = 30000$, $x_{min} = 250$.

Population 3: [Source: [3]].

First 12 observations are considered as a population.

y=Merchandise imports in dollars million,

x=Gross national product in dollars billion.

Data statistics are given as:

$N = 12$, $n = 5$, $\bar{Y} = 132.7323$, $\bar{X} = 1781.125$, $C_y = 0.5665$, $C_x = 0.3451$, $\rho_y^* = 0.2867$, $\rho_x^* = 0.2991$, $\rho_{yx} = 0.9954$, $y_{max} = 265.086$, $y_{min} = 39.866$, $x_{max} = 2957.8$, $x_{min} = 992.7$.

Population 4: [Source: [9]].

y= Output production of 40 factories.

x= The number of workers.

Data statistics are given as:

$N = 40$, $n = 10$, $\bar{Y} = 5078.575$, $\bar{X} = 230.3251$, $C_y = 0.329525$, $C_x = 0.84056$, $\rho_y^* = 0.23$, $\rho_x^* = 0.086$, $\rho_{yx} = 0.8005$, $y_{max} = 8512$, $y_{min} = 1451$, $x_{max} = 662$, $x_{min} = 52$.

The results are given in Table 1.

Table 1. MSE values of all considered estimators.

Estimator	Population 1	Population 2	Population 3	Population 4
$\bar{y}_{R(sys)}$	3.38671	0.01101	28.07364	40019.90
$\hat{\bar{Y}}_{R(sys)}$	3.37638	0.01094	28.03744	39976.72
$\bar{y}_{P(sys)}$	117.1967	0.08969	496.58950	291680.30
$\hat{\bar{Y}}_{P(sys)}$	117.1864	0.08962	496.55330	291637.10
$\bar{y}_{lr(sys)}$	1.40830	0.00997	1.73586	14634.74
$\hat{\bar{Y}}_{lr(g)(sys)}$	0.92753	0.00979	1.49143	14583.30

Table 1 shows the results based on four populations. For all four data sets, the recommended estimators' mean square error values are lower than the existing estimators. Among all the estimators investigated, the recommended regression estimator $\hat{\bar{Y}}_{lr(g)(sys)}$ outperforms and is preferred.

6. Conclusions

When employing maximum and minimum values, we provided certain ratio, product, and regression type estimators in a systematic sampling strategy. Under certain situations, the suggested estimators are expected to be more efficient than traditional ratio, product, and regression estimators. In all four populations, the performance of the recommended estimators is better than the standard estimators, as shown in Table 1. As a result, it's possible that the proposed estimators will be favored over the existing estimators. The recommended regression estimator has the best performance of all the estimators.

Acknowledgments

We would like to thank you for following the instructions above very closely in advance. It will definitely save us lot of time and expedite the process of your paper's publication.

Conflict of interest

There is no conflict of interest between the authors.

References

1. W. Gautschi, Some remarks on systematic sampling, *Ann. Math. Stat.*, **28** (1957), 385–394.
2. A. Griffith, The efficiency of enumerations, *Indian For.*, **73** (1947), 102–107.
3. D. N. Gujarati, *Basic econometrics*, Tata McGraw-Hill Education, 2003.
4. A. Hasel, Estimation of volume in timber stands by strip sampling, *Ann. Math. Stat.*, **13** (1942), 179–206.
5. K. Kushwaha, H. Singh, Class of almost unbiased ratio and product estimators in systematic sampling, *J. Ind. Soc. Agric. Stat.*, **41** (1989), 193–205.

6. S. N. S. Kushwaha, K. S. Kushwaha, A class of ratio, product and difference (rpd) estimators in systematic sampling, *Microelectron. Reliab.*, **33** (1993), 455–457. https://doi.org/10.1016/0026-2714(93)90308-L
7. D. Lahiri, On the question of bias of systematic sampling, *Proceedings of world population conference*, **6** (1954), 349–362.
8. C. Long, W. Chen, R. Yang, D. Yao, Ratio estimation of the population mean using auxiliary information under the optimal sampling design, *Probab. Eng. Inf. Sci.*, 2020, 1–12. https://doi.org/10.1017/S0269964820000625
9. P. Mukhopadhyay, *Theory and methods of survey sampling*, New Delhi: PHI Learning Pvivate Limited, 2 Eds., 2009.
10. C. E. Sarndal, Sample survey theory vs. general statistical theory: Estimation of the population mean, *Int. Stat. Rev.*, **40** (1972), 1–12. https://doi.org/10.2307/1402700
11. N. Shukla, Systematic sampling and product method of estimation, *Proceeding of all India Seminar on Demography and Statistics*, BHU, Varanasi, India, 1971.
12. D. Singh, F. S. Chaudhary, *Theory and analysis of sample survey designs*, John Wiley & Sons, 1986.
13. R. Singh, H. Singh, Almost unbiased ratio and product-type estimators in systematic sampling, *QÜESTIIÓ*, **22** (1998), 403–416.
14. A. Swain, The use of systematic sampling in ratio estimate, *J. Ind. Stat. Assoc.*, **2** (1964), 160–164.

AIMS Press