

Research Article

Finite Population Mean Estimation Under Systematic Sampling Scheme in Presence of Maximum and Minimum Values Using Two Auxiliary Variables

Mahmoud El-Morshedy ^{1,2}, Sardar Hussain ³, Kalim Ullah,⁴ Alamgir Khalil,⁵ Javid Shabbir ^{3,6} and Wahidullah Mansoor ⁷

¹Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

²Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

³Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

⁴Foundation Foundation University Medical College, Foundation University School of Health Sciences Islamabad, DHA-I Islamabad 44000, Pakistan

⁵Department of Statistics, University of Peshawar, Peshawar, Pakistan

⁶Department of Statistics, University of Wah, Wah Cantt, Pakistan

⁷Department of Biology, Yazd University, P. O. Box 89175-741, Yazd, Iran

Correspondence should be addressed to Wahidullah Mansoor; wahidmansoor@stu.yazd.ac.ir

Received 3 June 2022; Accepted 31 August 2022; Published 28 September 2022

Academic Editor: Tahir Mehmood

Copyright © 2022 Mahmoud El-Morshedy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The purpose of this paper is to look at how to estimate the finite population mean utilizing information from the auxiliary variable on a systematic sampling technique. By integrating the study variable's maximum and minimum values, as well as two auxiliary variables, we offer estimators of the ratio, product, and regression types. The mathematical equations of the suggested and existing estimators are derived up to the first order of approximation. Based on real-life data sets, efficiency comparisons are carried out. The suggested ratio, product and regression estimators consistently outperform existing estimators in terms of mean square error, according to theoretical and empirical investigations.

1. Introduction

It is common practise in sampling surveys to employ auxiliary data during the estimation stage to improve the precision of population parameter estimates. In this context, several ratios, products, and regression estimators are appropriate examples. In the literature, a larger number of consistent estimators using auxiliary information for estimating the finite population mean or a total of the study variable, as well as their properties in simple random sampling, have been addressed, for example, see [1, 2] and the references cited therein. For estimating the population characteristics, there are some populations, such as forest

areas for estimating timber volume and areas under various forms of cover [3], where simple random sampling or other sampling schemes are difficult to apply. In such a setting, systematic sampling produces precise results for selecting a sample from a population. The advantages of systematic sampling include the ability to select the entire sample with just one random start.

Also sampling theory, appropriate use of the auxiliary information may increase the precision of the estimators. But unfortunately, many real data sets contain values that are suddenly maximum or minimum. As a result, if any unexpected values are chosen in the sample, the estimator may generate misleading results. To handle such situations, we

proposed mean estimation under a systematic sampling design in the presence of maximum and minimum values.

Assume that the population units N are sequentially numbered from 1 to N . We choose a unit at random from the first k units and every k_{th} unit to create a sample of n units. As an example, if k is 12 and the first unit drawn is number 10, the following units are numbers 22, 34, 44, and so on.

The whole number is determined by the first unit chosen. A simple random sample appears to be less exact than systematic sampling. It divides the population into n strata, each of which includes the first k units, second k units, and so on. References [4, 5] discovered that systematic sampling is efficient and convenient in sampling. It gives estimators that are more efficient than those offered by basic random sampling under certain actual situations, in addition to its simplicity, which is very important in large-scale sampling work. The ratio and product estimators for estimating the finite population mean \bar{y} of the study variable y were built by [6, 7]. References [7–15] and Javaid et al. [16], all go into great length about systematic sampling.

2. Notation and Symbols

Consider the study variable Y and the auxiliary variables X, Z for a finite population with $U = \{U_1, U_2, \dots, U_N\}$ units. We choose a systemic sample of size n , starting with a random selection of the first unit, and subsequently selecting every k th unit after each interval of k . We'll use the formula $N = nk$, with n and k being positive integers. For selected systematic random sample, say y_{ij}, x_{ij} , where $i = 1, 2, \dots, n, j = 1, 2, \dots, k$, are the values of j^{th} unit in the i^{th} selected sample for Y, X , and Z variables correspondingly. For y, x , and z , the sample mean in systematic random sampling is

$$\begin{aligned}\widehat{\bar{y}}_{sys} &= \frac{1}{n} \sum_{j=1}^n y_{ij}, \\ \widehat{\bar{x}}_{sys} &= \frac{1}{n} \sum_{j=1}^n x_{ij}, \\ \widehat{\bar{z}}_{sys} &= \frac{1}{n} \sum_{j=1}^n z_{ij},\end{aligned}\quad (1)$$

those are unbiased estimators for population means for the variables Y, X , and Z , respectively. The bias and mean squared error were calculated, let us define

$$\begin{aligned}e_0 &= \frac{\widehat{\bar{y}}_{sys} - \bar{y}}{\bar{y}}, \\ e_1 &= \frac{\widehat{\bar{x}}_{sys} - \bar{x}}{\bar{x}}, \\ e_2 &= \frac{\widehat{\bar{z}}_{sys} - \bar{z}}{\bar{z}},\end{aligned}\quad (2)$$

such that, $E(e_0) = E(e_1) = E(e_2) = 0$

$$\begin{aligned}E(e_0^2) &= \lambda \rho_y^{*2} C_y^2, \\ E(e_1^2) &= \lambda \rho_x^{*2} C_x^2, \\ E(e_2^2) &= \lambda \rho_z^{*2} C_z^2, \\ E(e_0 e_1) &= \lambda \rho_{yx} C_y C_x \sqrt{\rho_y^* \rho_x^*}, \\ E(e_0 e_2) &= \lambda \rho_{yz} C_y C_z \sqrt{\rho_y^* \rho_z^*}, \\ E(e_1 e_2) &= \lambda \rho_{xz} C_x C_z \sqrt{\rho_x^* \rho_z^*},\end{aligned}\quad (3)$$

where

$$\begin{aligned}\rho_y^* &= 1 + (n-1)\rho_y, \\ \rho_x^* &= 1 + (n-1)\rho_x, \\ \rho_z^* &= 1 + (n-1)\rho_z, \\ C_y &= \frac{S_y}{\bar{y}}, \\ S_y &= \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2}, \\ C_x &= \frac{S_x}{\bar{x}}, \\ S_x &= \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}, \\ C_z &= \frac{S_z}{\bar{z}}, \\ S_z &= \sqrt{\frac{1}{N-1} \sum_{i=1}^N (z_i - \bar{z})^2}.\end{aligned}\quad (4)$$

The coefficients of y , x , and z are C_y , C_x , and C_z , respectively, and

$$\begin{aligned} \rho_y &= \frac{E\left[(y_{ij} - \bar{y})(y_{ij} - \bar{y})\right]}{E(y_{ij} - \bar{y})}, \\ \rho_x &= \frac{E\left[(x_{ij} - \bar{x})(x_{ij} - \bar{x})\right]}{E(x_{ij} - \bar{x})}, \\ \rho_z &= \frac{E\left[(z_{ij} - \bar{z})(z_{ij} - \bar{z})\right]}{E(z_{ij} - \bar{z})}, \end{aligned} \quad (5)$$

are interclass correlation in the systematic sample for the research variable y as well as the auxiliary variables x and z ,

$$\begin{aligned} \rho_{yx} &= \frac{E(y_{ij} - \bar{y})(x_{ij} - \bar{x})}{\sqrt{E(y_{ij} - \bar{y})^2 E(x_{ij} - \bar{x})^2}}, \\ \rho_{xz} &= \frac{E(x_{ij} - \bar{x})(z_{ij} - \bar{z})}{\sqrt{E(x_{ij} - \bar{x})^2 E(z_{ij} - \bar{z})^2}}, \\ \rho_{yz} &= \frac{E(y_{ij} - \bar{y})(z_{ij} - \bar{z})}{\sqrt{E(y_{ij} - \bar{y})^2 E(z_{ij} - \bar{z})^2}}, \end{aligned} \quad (6)$$

is the correlation between the study variable y and auxiliary variables (x, z).

Many genuine data sets have unexpectedly large or tiny (y_{\max}) or (y_{\min}) values. When such values appear in the estimation of a finite population mean, the results are vulnerable. The findings will be either inflated or underestimated if y_{\max} and y_{\min} exist. To deal with such a situation, [17] suggested the following unbiased estimator for the estimation of finite population mean using maximum and minimum values:

$$\% \hat{\bar{y}}_s = \begin{cases} \hat{\bar{y}} + c, & \text{if samples contain } y_{\min} \text{ but not } y_{\max} \\ \hat{\bar{y}} - c, & \text{if samples contain } y_{\max} \text{ but not } y_{\min} \\ \hat{\bar{y}}, & \text{for all other samples.} \end{cases} \quad (7)$$

The variance of $\hat{\bar{y}}_s$, is calculated as follows:

$$V\left(\hat{\bar{y}}_s\right) = \lambda S_y^2 - \frac{2\lambda nc}{N-1} (y_{\max} - y_{\min} - nc), \quad (8)$$

the population variance is S_y^2 , and the constant is c . c has a minimum value of

$$c_{\text{opt}} = \frac{(y_{\max} - y_{\min})^2}{2n}. \quad (9)$$

The variance of $\hat{\bar{y}}_s$, is calculated as follows:

$$V\left(\hat{\bar{y}}_s\right)_{\min} = V\left(\hat{\bar{y}}\right) - \frac{\lambda(y_{\max} - y_{\min})^2}{2(N-1)}, \quad (10)$$

this is always less than the variance of $\hat{\bar{y}}$.

Under a systematic sampling approach, the usual ratio estimator is

$$\hat{\bar{y}}_{\text{Rsys}} = \hat{\bar{y}}_{\text{sys}} \left(\frac{\bar{x}}{\hat{\bar{x}}_{\text{sys}}} \right) \left(\frac{\bar{z}}{\hat{\bar{z}}_{\text{sys}}} \right). \quad (11)$$

The bias and mean squared error of $\hat{\bar{y}}_{\text{Rsys}}$ are provided by the following up to the first order of approximation:

$$\begin{aligned} \text{Bias}\left(\hat{\bar{y}}_{\text{sys}}\right) &= \bar{y}\lambda \left[C_x^2 \rho_x^* + C_z^2 \rho_z^* - \rho_{yx} C_y C_x \sqrt{\rho_y^* \rho_x^*} \right. \\ &\quad \left. - \rho_{yz} C_y C_z \sqrt{\rho_y^* \rho_z^*} + \rho_{xz} C_x C_z \sqrt{\rho_x^* \rho_z^*} \right], \end{aligned} \quad (12)$$

and

$$\begin{aligned} \text{MSE}\left(\hat{\bar{y}}_{\text{Rsys}}\right) &= \lambda \left[S_y^2 \rho_y^* + R_1^2 S_x^2 \rho_x^* + R_2^2 S_z^2 \rho_z^* - 2R_1 S_{yx} \sqrt{\rho_y^* \rho_x^*} \right. \\ &\quad \left. + 2R_1 R_2 S_{xz} \sqrt{\rho_x^* \rho_z^*} - 2R_2 S_{yz} \sqrt{\rho_y^* \rho_z^*} \right], \end{aligned} \quad (13)$$

where, $R_1 = \bar{y}/\bar{x}$ and $R_2 = \bar{y}/\bar{z}$ Under systematic sampling, the product estimator is as follows:

$$\hat{\bar{y}}_{P(\text{sys})} = \hat{\bar{y}}_{\text{sys}} \left(\frac{\hat{\bar{x}}_{\text{sys}}}{\bar{x}} \right) \left(\frac{\hat{\bar{z}}_{\text{sys}}}{\bar{z}} \right). \quad (14)$$

The bias and MSE of $\hat{\bar{y}}_{P(\text{sys})}$ are provided by the following up to the first order of approximation:

$$\begin{aligned} \text{Bias}\left(\hat{\bar{y}}_{P(\text{sys})}\right) &= \bar{y}\lambda \left[\rho_{yx} C_y C_x \sqrt{\rho_y^* \rho_x^*} + \rho_{yz} C_y C_z \sqrt{\rho_y^* \rho_z^*} \right. \\ &\quad \left. + \rho_{xz} C_x C_z \sqrt{\rho_x^* \rho_z^*} \right], \end{aligned} \quad (15)$$

and

$$\begin{aligned} \text{MSE}\left(\hat{\bar{y}}_{P(\text{sys})}\right) &= \lambda \left[S_y^2 \rho_y^* + R_1^2 S_x^2 \rho_x^* + R_2^2 S_z^2 \rho_z^* - 2R_1 S_{yx} \sqrt{\rho_y^* \rho_x^*} \right. \\ &\quad \left. + 2R_1 R_2 S_{xz} \sqrt{\rho_x^* \rho_z^*} - 2R_2 S_{yz} \sqrt{\rho_y^* \rho_z^*} \right]. \end{aligned} \quad (16)$$

Under a systematic sampling procedure, the standard regression estimator for predicting the unknown population mean is as follows:

$$\hat{\bar{y}}_{lr(\text{sys})} = \hat{\bar{y}}_{\text{sys}} + b_1 (\bar{x} - \hat{\bar{x}}_{\text{sys}}) + b_2 (\bar{z} - \hat{\bar{z}}_{\text{sys}}). \quad (17)$$

The sample regression coefficients are b_1 and b_2 , respectively. If b_1 and b_2 are the least square estimators of B_1 and B_2 , respectively, then up to the first order of approximation, the variance of the estimator $\hat{\bar{y}}_{lr(\text{sys})}$ is as follows:

$$V(\widehat{\mathcal{Y}}_{lr(\text{sys})}) = \lambda S_y^2 \rho_y^* [1 - (\rho_{yx}^2 + \rho_{yz}^2 - 2\rho_{yx}\rho_{yz}\rho_{xz})]. \quad (18)$$

3. Suggested Estimators

We provide a ratio, product, and regression type estimator using auxiliary variables and the study variable in a systematic sampling method, based on [18]. We also take into account the study's minimum and maximum values, as well as the two auxiliary variables.

3.1. First Situation. When the study variable and the auxiliary variable have a positive connection, a bigger value of

the auxiliary variable should be chosen, as should a greater value of the study variable. A smaller study variable value should also be chosen, as well as a smaller value of the auxiliary variable. To make use of these type of data, using auxiliary variables and the research variable, we recommend using a ratio type estimator.

$$\widehat{\mathcal{Y}}_{R_{\text{suggested}(\text{sys})}} = \widehat{\mathcal{Y}}_{\text{sys}} \left(\frac{\overline{\mathcal{X}}}{\widehat{\mathcal{X}}_{\text{sys}}} \right) \left(\frac{\overline{\mathcal{Z}}}{\widehat{\mathcal{Z}}_{\text{sys}}} \right). \quad (19)$$

or

$$\widehat{\mathcal{Y}}_{R_{\text{suggested}(\text{sys})}} \begin{cases} \left(\widehat{\mathcal{Y}}_{\text{sys}} + c_1 \right) \left(\frac{\overline{\mathcal{X}}}{\widehat{\mathcal{X}}_{\text{sys}} + c_2} \right) \left(\frac{\overline{\mathcal{Z}}}{\widehat{\mathcal{Z}}_{\text{sys}} + c_3} \right), & \text{if sample contains } y_{\min} \text{ and } x_{\min}, z_{\min} \\ \left(\widehat{\mathcal{Y}}_{\text{sys}} - c_1 \right) \left(\frac{\overline{\mathcal{X}}}{\widehat{\mathcal{X}}_{\text{sys}} - c_2} \right) \left(\frac{\overline{\mathcal{Z}}}{\widehat{\mathcal{Z}}_{\text{sys}} - c_3} \right), & \text{if sample contains } y_{\max} \text{ and } x_{\max}, z_{\max} \\ \widehat{\mathcal{Y}}_{\text{sys}} \left(\frac{\overline{\mathcal{X}}}{\widehat{\mathcal{X}}_{\text{sys}}} \right) \left(\frac{\overline{\mathcal{Z}}}{\widehat{\mathcal{Z}}_{\text{sys}}} \right), & \text{for all the rest the samples} \end{cases} \quad (20)$$

The estimator of regression type is:

$$\widehat{\mathcal{Y}}_{lr_{\text{suggested}(\text{sys})}} = \widehat{\mathcal{Y}}_{(\text{sys})c_{11}} + b_1 \left(\overline{\mathcal{X}} - \widehat{\mathcal{X}}_{(\text{sys})c_{21}} \right) + b_2 \left(\overline{\mathcal{Z}} - \widehat{\mathcal{Z}}_{(\text{sys})c_{31}} \right), \quad (21)$$

where, $\widehat{\mathcal{Y}}_{c_{11}} = \widehat{\mathcal{Y}} + c_1$, $\widehat{\mathcal{X}}_{c_{21}} = \widehat{\mathcal{X}} + c_2$, $\widehat{\mathcal{Z}}_{c_{31}} = \widehat{\mathcal{Z}} + c_3$. If the sample contains y_{\min} and (x_{\min}, z_{\min}) , then $\widehat{\mathcal{Y}}_{c_{21}} = \widehat{\mathcal{Y}} - c_1$, $\widehat{\mathcal{X}}_{c_{22}} = \widehat{\mathcal{X}} - c_2$, $\widehat{\mathcal{Z}}_{c_{32}} = \widehat{\mathcal{Z}} - c_3$. If the samples contain y_{\max} and (x_{\max}, z_{\max}) , then $\widehat{\mathcal{Y}}_{c_{11}} = \widehat{\mathcal{Y}}$, $\widehat{\mathcal{X}}_{c_{21}} = \widehat{\mathcal{X}}$, $\widehat{\mathcal{Z}}_{c_{31}} = \widehat{\mathcal{Z}}$, for all other samples.

3.2. Second Situation. When the study variable and the auxiliary variable have a negative correlation, choose the auxiliary variable with the bigger value and the study variable with the smallest value. The smaller value of the auxiliary variable is chosen, while the larger value of the study variable is chosen. The proposed product type estimator using the auxiliary variables (x, z) with the study variable y is given by

$$\widehat{\mathcal{Y}}_{P_{\text{suggested}(\text{sys})}} = \widehat{\mathcal{Y}}_{(\text{sys})c_{12}} \left(\frac{\widehat{\mathcal{X}}_{(\text{sys})c_{22}}}{\overline{\mathcal{X}}} \right) \left(\frac{\widehat{\mathcal{Z}}_{(\text{sys})c_{32}}}{\overline{\mathcal{Z}}} \right). \quad (22)$$

or

$$\widehat{\mathcal{Y}}_{P_{\text{suggested}}(\text{sys})} = \begin{cases} \left(\widehat{\mathcal{Y}}_{\text{sys}} + c_1 \right) \left(\frac{\widehat{\mathcal{X}}_{\text{sys}} + c_2}{\widehat{\mathcal{X}}} \right) \left(\frac{\widehat{\mathcal{Z}}_{\text{sys}} + c_3}{\widehat{\mathcal{Z}}} \right), & \text{if samples contain } y_{\min} \text{ and } x_{\max}, z_{\max} \\ \left(\widehat{\mathcal{Y}}_{\text{sys}} + c_1 \right) \left(\frac{\widehat{\mathcal{X}}_{\text{sys}} + c_2}{\widehat{\mathcal{X}}} \right) \left(\frac{\widehat{\mathcal{Z}}_{\text{sys}} + c_3}{\widehat{\mathcal{Z}}} \right), & \text{if samples contain } y_{\max} \text{ and } x_{\min}, z_{\min} \\ \widehat{\mathcal{Y}}_{\text{sys}} \left(\frac{\widehat{\mathcal{X}}_{\text{sys}}}{\widehat{\mathcal{X}}} \right) \left(\frac{\widehat{\mathcal{Z}}_{\text{sys}}}{\widehat{\mathcal{Z}}} \right), & \text{for all other samples.} \end{cases} \quad (23)$$

The estimator of regression type is as follows:

$$\widehat{\mathcal{Y}}_{R_{\text{suggested}}(\text{sys})} = \widehat{\mathcal{Y}}_{(\text{sys})c_{11}} + b_1 \left(\widehat{\mathcal{X}} - \widehat{\mathcal{X}}_{(\text{sys})c_{22}} \right) + b_2 \left(\widehat{\mathcal{Z}} - \widehat{\mathcal{Z}}_{(\text{sys})c_{32}} \right), \quad (24)$$

where $(\widehat{\mathcal{Y}}_{c_{12}} = \widehat{\mathcal{Y}} + c_1, \widehat{\mathcal{X}}_{c_{22}} = \widehat{\mathcal{X}} - c_2, \widehat{\mathcal{Z}}_{c_{32}} = \widehat{\mathcal{Z}} + c_3)$. If the samples contain y_{\min} and (x_{\max}, z_{\max}) , and also $(\widehat{\mathcal{Y}}_{c_{11}} = \widehat{\mathcal{Y}} - c_1, \widehat{\mathcal{X}}_{c_{21}} = \widehat{\mathcal{X}} + c_2, \widehat{\mathcal{Z}}_{c_{32}} = \widehat{\mathcal{Z}} + c_3)$. If the samples contain y_{\max} and (x_{\min}, z_{\min}) and $(\widehat{\mathcal{Y}}_{c_{12}} = \widehat{\mathcal{Y}}, \widehat{\mathcal{X}}_{c_{22}} = \widehat{\mathcal{X}}, \widehat{\mathcal{Z}}_{c_{32}} = \widehat{\mathcal{Z}})$ for all types of samples. Also $c_1, c_2,$ and c_3 are unknown constants. The following relative error terms and their expectations are used to generate biases and mean squared errors.

$$\begin{aligned} e_0 &= \frac{\widehat{\mathcal{Y}}_{(\text{sys})c_1} - \overline{\mathcal{Y}}}{\overline{\mathcal{Y}}}, \\ e_1 &= \frac{\widehat{\mathcal{X}}_{(\text{sys})c_2} - \overline{\mathcal{X}}}{\overline{\mathcal{X}}}, \\ e_2 &= \frac{\widehat{\mathcal{Z}}_{(\text{sys})c_3} - \overline{\mathcal{Z}}}{\overline{\mathcal{Z}}}, \end{aligned} \quad (25)$$

such that

$$E(e_0) = E(e_1) = E(e_2) = 0$$

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

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

$$E(e_2^2) = \frac{\lambda}{\overline{\mathcal{Z}}^2} \left[S_z^2 \rho_z^* - \frac{2nc_3}{N-1} (\Delta_z - nc_3) \right],$$

$$E(e_0 e_1) = \frac{\lambda}{\overline{\mathcal{Y}} \overline{\mathcal{X}}} \left[S_{yx} \sqrt{\rho_y^* \rho_x^*} - \frac{n}{N-1} \{c_2 \Delta_y + c_1 \Delta_x - 2nc_1 c_2\} \right],$$

$$E(e_0 e_2) = \frac{\lambda}{\overline{\mathcal{Y}} \overline{\mathcal{Z}}} \left[S_{yz} \sqrt{\rho_y^* \rho_z^*} - \frac{n}{N-1} \{c_3 \Delta_y + c_1 \Delta_z - 2nc_1 c_3\} \right],$$

$$E(e_1 e_2) = \frac{\lambda}{\overline{\mathcal{X}} \overline{\mathcal{Z}}} \left[S_{xz} \sqrt{\rho_x^* \rho_z^*} - \frac{n}{N-1} \{c_3 \Delta_x + c_2 \Delta_z - 2nc_2 c_3\} \right]. \quad (26)$$

Expressing (19) in terms of e 's, we have

$$\widehat{\mathcal{Y}}_{R_{\text{suggested}}(\text{sys})} = \overline{\mathcal{Y}} (1 + e_0) (1 + e_1)^{-1} (1 + e_2)^{-1}. \quad (27)$$

Expressing (27) we have the following results up to the first order of approximation:

$$\widehat{\mathcal{Y}}_{R_{\text{suggested}}(\text{sys})} - \overline{\mathcal{Y}} = \overline{\mathcal{Y}} (e_0 - e_1 - e_2 + e_2^2 + e_1^2 + e_1 e_2 - e_0 e_2 - e_0 e_1). \quad (28)$$

Taking both sides of (28) into consideration, we have

$$\text{Bias} \left(\widehat{\mathcal{Y}}_{R_{\text{suggested}}(\text{sys})} \right) \cong \lambda \begin{bmatrix} \frac{R_1}{\overline{\mathcal{X}} \overline{\mathcal{Z}}} \left\{ S_x^2 \rho_x^* - \frac{2nc_2}{N-1} (\Delta_x - nc_2) \right\} + R_2 \left\{ S_z^2 \rho_z^* - \frac{2nc_3}{N-1} (\Delta_z - nc_3) \right\} \\ -R_1 R_2 \left\{ S_{xz} \sqrt{\rho_x^* \rho_z^*} - \frac{n}{N-1} \{c_3 \Delta_x + c_2 \Delta_z - 2nc_2 c_3\} \right\} \\ -R_1 \left\{ S_{yx} \sqrt{\rho_y^* \rho_x^*} - \frac{n}{N-1} \{c_2 \Delta_y + c_1 \Delta_x - 2nc_1 c_2\} \right\} \\ -R_2 \left\{ S_{yz} \sqrt{\rho_y^* \rho_z^*} - \frac{n}{N-1} \{c_3 \Delta_y + c_2 \Delta_z - 2nc_2 c_3\} \right\} \end{bmatrix}. \quad (29)$$

We have after squaring (28) and taking expectations

$$\text{MSE}\left(\widehat{\mathcal{Y}}_{R_{\text{suggested}(\text{sys})}}\right) \cong \left[\begin{array}{l} \lambda \left(S_y^2 \rho_y^* + R_1^2 S_x^2 \rho_x^* + R_2^2 S_z^2 \rho_z^* + 2R_1 R_2 S_{yx} \sqrt{\rho_y^* \rho_x^*} - 2R_2 S_{yz} \sqrt{\rho_y^* \rho_z^*} \right) \\ - \frac{2n\lambda (c_1 R_2 c_3 - R_1 c_2)}{N-1} \{ \Delta_y - R_1 \Delta_x - R_2 \Delta_z - n(c_1 - R_2 c_3 - R_1 c_2) \} \end{array} \right]. \quad (30)$$

Differentiate (30) we have, c_1 , c_2 , and c_3

$$\begin{aligned} c_{1\text{opt}} &= \frac{y_{\max} - y_{\min}}{2n}, \\ c_{2\text{opt}} &= \frac{x_{\max} - x_{\min}}{2n}, \\ c_{3\text{opt}} &= \frac{z_{\max} - z_{\min}}{2n}. \end{aligned} \quad (31)$$

We derive the minimum MSE by substituting the optimum values of c_1 , c_2 , and c_3 in (30) MSE of $(\widehat{\mathcal{Y}}_{R(\text{sys})})$, given by

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{R_{\text{suggested}(\text{sys})}}\right)_{\min} &\cong \text{MSE}\left(\widehat{\mathcal{Y}}_{R(\text{sys})}\right) \\ &= \frac{\lambda}{2(N-1)} [\Delta_y - R_1 \Delta_x - R_2 \Delta_z]^2, \end{aligned} \quad (32)$$

where

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{R_{\text{suggested}(\text{sys})}}\right) &= \lambda \left[\begin{array}{l} S_y^2 \rho_y^* + R_1^2 S_x^2 \rho_x^* + R_2^2 S_z^2 \rho_z^* - 2R_1 S_{yx} \sqrt{\rho_y^* \rho_x^*} \\ + 2R_1 R_2 S_{xz} \sqrt{\rho_x^* \rho_z^*} - 2R_2 S_{yz} \sqrt{\rho_y^* \rho_z^*} \end{array} \right]. \end{aligned} \quad (33)$$

Similarly

$$\text{Bias}\left(\widehat{\mathcal{Y}}_{P_{\text{suggested}(\text{sys})}}\right) \cong \lambda \left[\begin{array}{l} \frac{R_1}{\mathcal{X}\mathcal{Z}} \left\{ S_x^2 \rho_x^* - \frac{2nc_2}{N-1} \Delta_x - nc_2 \right\} + R_2 \left\{ S_{v^*}^2 - \frac{2nc_3}{N-1} \Delta_z - nc_3 \right\} \\ + R_1 R_2 \left\{ S_{yz} \sqrt{\rho_y^* \rho_z^*} - \frac{n}{N-1} \{ c_3 \Delta_x + c_2 \Delta_z - 2nc_2 c_3 \} \right\} \\ + R_1 \left\{ S_{yx} \sqrt{\rho_y^* \rho_x^*} - \frac{n}{N-1} \{ c_2 \Delta_y + c_1 \Delta_x - 2nc_1 c_2 \} \right\} \\ + R_2 \left\{ S_{uv^*} - \frac{n}{N-1} \{ c_3 \Delta_y + c_2 \Delta_z - 2nc_2 c_3 \} \right\} \end{array} \right], \quad (34)$$

and

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{P_{\text{suggested}(\text{sys})}}\right)_{\min} &\cong \text{MSE}\left(\widehat{\mathcal{Y}}_{P_{\text{suggested}(\text{sys})}}\right) \\ &= \frac{\lambda}{2(N-1)} [\Delta_y + R_1 \Delta_x + R_2 \Delta_z]^2, \end{aligned} \quad (35)$$

where

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{P_{\text{suggested}(\text{sys})}}\right) &= \lambda \left[\begin{array}{l} S_y^2 \rho_y^* + R_1^2 S_x^2 \rho_x^* + R_2^2 S_z^2 \rho_z^* - 2R_1 S_{yx} \sqrt{\rho_y^* \rho_x^*} \\ + 2R_1 R_2 S_{xz} \sqrt{\rho_x^* \rho_z^*} - 2R_2 S_{yz} \sqrt{\rho_y^* \rho_z^*} \end{array} \right]. \end{aligned} \quad (36)$$

In the case of positive correlation, the minimum MSE of the regression estimator is provided by

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right)_{\min} &\cong \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right) \\ &- \frac{\lambda}{2(N-1)}[\Delta_y - \beta_1\Delta_x - \beta_2\Delta_z]^2, \end{aligned} \quad (37)$$

where

$$\begin{aligned} V\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right) &= \lambda S_y^2 \rho_y^* [1 - (\rho_{yx}^2 + \rho_{yz}^2 - 2\rho_{yx}\rho_{yz}\rho_{xz})], \\ \beta_1 &= \rho_{yx} \frac{S_y \sqrt{\rho_y^*}}{S_x \sqrt{\rho_x^*}}, \\ \beta_2 &= \rho_{yz} \frac{S_y \sqrt{\rho_y^*}}{S_z \sqrt{\rho_z^*}}. \end{aligned} \quad (38)$$

The population regression coefficients are β_1 and β_2 .

Similarly, in the case of negative correlation, the minimal MSE of $\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}$ is

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right)_{\min} &\cong \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right) \\ &- \frac{\lambda}{2(N-1)}[\Delta_y + \beta_1\Delta_x + \beta_2\Delta_z]^2. \end{aligned} \quad (39)$$

In the case of both positive and negative correlation between the study and the auxiliary variable, a general form for MSE is

$$\begin{aligned} \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right)_{\min} &\cong \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right) \\ &- \frac{\lambda}{2(N-1)}[\Delta_y - |\beta_1|\Delta_x - |\beta_2|\Delta_z]^2. \end{aligned} \quad (40)$$

4. Comparison of Estimators

In this part, we use a systematic sampling approach to compare the suggested estimators against classic ratios, products, and regression estimators.

Condition (i).

By (13) and (32)

$$\left[\text{MSE}\left(\widehat{\mathcal{Y}}_{R(\text{sys})}\right) - \text{MSE}\left(\widehat{\mathcal{Y}}_{R_{\text{suggested (sys)}}}\right)_{\min} \right] \geq 0, \quad (41)$$

if

$$[\Delta_y - R_1\Delta_x - R_2\Delta_z]^2 \geq 0. \quad (42)$$

Condition (ii).

By (16) and (33)

TABLE 1: MSE.

Estimator	Data set 1	Data set 2
$\widehat{\mathcal{Y}}_{R(\text{sys})}$	317.2277	307097.1
$\widehat{\mathcal{Y}}_{R_{\text{suggested (sys)}}$	309.0105	256051.6
$\widehat{\mathcal{Y}}_{P(\text{sys})}$	317.2294	307097.4
$\widehat{\mathcal{Y}}_{P_{\text{suggested (sys)}}$	309.0122	256051.6
$\widehat{\mathcal{Y}}_{lr(\text{sys})}$	187.2088	47065.9
$\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}$	154.5798	18097.1

TABLE 2: PRE.

Estimator	Data set 1	Data set 2
$\widehat{\mathcal{Y}}_{R(\text{sys})}$	100.0000	100.0000
$\widehat{\mathcal{Y}}_{R_{\text{suggested (sys)}}$	102.6599	119.9356
$\widehat{\mathcal{Y}}_{P(\text{sys})}$	100.0000	100.0000
$\widehat{\mathcal{Y}}_{P_{\text{suggested (sys)}}$	102.6592	119.9357
$\widehat{\mathcal{Y}}_{lr(\text{sys})}$	100.0000	100.0000
$\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}$	121.1080	260.0743

$$\left[\text{MSE}\left(\widehat{\mathcal{Y}}_{P(\text{sys})}\right) - \text{MSE}\left(\widehat{\mathcal{Y}}_{P_{\text{suggested (sys)}}}\right)_{\min} \right] \geq 0, \quad (43)$$

if

$$[\Delta_y + R_1\Delta_x + R_2\Delta_z]^2 \geq 0. \quad (44)$$

Condition (iii).

By (18) and (37)

$$\left[\text{MSE}\left(\widehat{\mathcal{Y}}_{lr(\text{sys})}\right) - \text{MSE}\left(\widehat{\mathcal{Y}}_{lr_{\text{suggested (sys)}}}\right)_{\min} \right] \geq 0, \quad (45)$$

if

$$[\Delta_y - |\beta_1|\Delta_x - |\beta_2|\Delta_z]^2 \geq 0. \quad (46)$$

We found that if Conditions (i)–(iii) are met, the suggested estimators outperform the existing estimators.

5. Empirical Study

For numerical comparisons, we utilize two distinct data sets:

Population 1: (Source: [19]).

y = Imports of merchandise in millions.

x = In billions, the gross national GDP.

z = Cumulative price in index, all items (1967 = 100).

$N = 12$, $n = 5$,

$\bar{y} = 132.7322$, $\bar{x} = 1781.125$, $\bar{z} = 174.075$,

$C_y = 0.5665$, $C_x = 0.3451$, $C_z = 0.2794$, $S_y = 75.19282$,

$S_x = 614.6662$, $S_z = 48.63655$, $\rho_{yx} = 0.9949$,

$$\rho_{yz} = 0.9916, \rho_{xz} = 0.9949, \rho_y = 0.2867, \rho_x = 0.2991, \\ \rho_z = 0.3421, S_{yx} = 0.08535, S_{xz} = 0.1018, S_{yz} = 0.0972, \\ y_{\max} = 265.086, y_{\min} = 39.866, x_{\max} = 2957.8, \\ x_{\min} = 992.7, z_{\max} = 272.41, z_{\min} = 116.3.$$

Population 2: (Source: [20])

y = Production output from 40 factories.

x = The total number of employees.

z = Fixed capital.

$N = 40, n = 10,$

$$\bar{y} = 5078.575, \bar{x} = 230.325, \bar{z} = 945.425, C_y = \\ 0.329525, C_x = 0.84056, C_z = 0.67555, S_y = 1673.517, \\ S_x = 193.602, S_z = 638.6819, \rho_{yx} = 0.8005, \rho_{yz} = \\ 0.83491, \rho_{xz} = 0.98161, \rho_y = 0.23, \rho_x = 0.086, \\ \rho_z = 0.16, y_{\max} = 8512, y_{\min} = 1451, x_{\max} = 662, \\ x_{\min} = 52, z_{\max} = 2355, z_{\min} = 107.$$

From the results of MSEs, which are available in Table 1, the mean squared errors of the suggested estimators are lower than the existing estimators. We can also see from Table 2, that PREs of all suggested estimators are surpassing all the existing estimators. Of all the suggested estimators, the regression estimator has the best performance.

6. Conclusion

We proposed several standard ratios, products, and regression estimators using auxiliary variables in a systematic sampling method in the presence of maximum and minimum values. Under some conditions, the proposed estimators are more efficient than traditional mean ratios, products, and regression estimators. Table 1 shows that the suggested estimators outperform the standard estimators in both populations. The numerical study also supports the superiority of our proposed estimators. It is found that the new suggested estimators of the finite population mean are more precise than some of the existing estimators.

Data Availability

The data used to support the numerical findings of this study are available from the corresponding author upon request. The data can also be obtained upon searching the given sources of data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was fully supported by the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan. The first author will pay the fee for this paper.

References

- [1] W. G. Cochran, "The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce," *The Journal of Agricultural Science*, vol. 30, no. 2, pp. 262–275, 1940.
- [2] P. V. Sukhatme, *Sampling Theory of Surveys with Applications*, The Indian Society Of Agricultural Statistics, New Delhi, 1957.
- [3] J. G. Osborne, "Sampling errors of systematic and random surveys of cover-type areas," *Journal of the American Statistical Association*, vol. 37, no. 218, pp. 256–264, 1942.
- [4] A. A. Hasel, "Estimation of volume in timber stands by strip sampling," *The Annals of Mathematical Statistics*, vol. 13, no. 2, pp. 179–206, 1942.
- [5] A. Griffith, "The efficiency of enumerations," *Indian Forester*, vol. 73, no. 3, pp. 102–107, 1947.
- [6] A. Swain, "The use of systematic sampling in ratio estimate," *Jour. Ind. Stat. Assoc.*, vol. 2, no. 213, pp. 160–164, 1964.
- [7] N. Shukla, "Systematic sampling and product method of estimation," in *Proceedings of the All India Seminar on Demography and Statistics*, BHU, Varanasi, India, 1971.
- [8] D. Lahiri, "On the question of bias of systematic sampling," *Proceedings - World Petroleum Congress*, vol. 6, pp. 349–362, 1954.
- [9] W. Gautschi, "Some remarks on systematic sampling," *The Annals of Mathematical Statistics*, vol. 28, no. 2, pp. 385–394, 1957.
- [10] K. Kushwaha and H. Singh, "Class of almost unbiased ratio and product estimators in systematic sampling," *Jour. Ind. Soc. Ag. Statistics*, vol. 41, no. 2, pp. 193–205, 1989.
- [11] P. K. Bedi, "Efficient utilization of auxiliary information at estimation stage," *Biometrical Journal*, vol. 38, no. 8, pp. 973–976, 1996.
- [12] S. Kushwaha, S. Kushwaha, and K. Kushwaha, "A class of ratio, product and difference (rpd) estimators in systematic sampling," *Microelectronics Reliability*, vol. 33, no. 4, pp. 455–457, 1993.
- [13] G. N. Singh and M. Khalid, "Exponential chain dual to ratio and regression type estimators of population mean in two-phase sampling," *Statistica*, vol. 75, no. 4, pp. 379–389, 2015.
- [14] G. N. Singh and M. Khalid, "Efficient class of estimators for finite population mean using auxiliary information in two-occasion successive sampling," *Journal of Modern Applied Statistical Methods*, vol. 17, no. 2, p. 14, 2019.
- [15] M. Noor-ul Amin, A. Javaid, and M. Hanif, "Estimation of population mean in systematic random sampling using auxiliary information," *Journal of Statistics & Management Systems*, vol. 20, no. 6, pp. 1095–1106, 2017.
- [16] A. Javaid, M. Noor-ul Amin, and M. Hanif, "Modified ratio estimator in systematic random sampling under non-response," *Proceedings of the National Academy of Sciences, India, Section A: Physical Sciences*, vol. 89, no. 4, pp. 817–825, 2019.
- [17] C.-E. Särndal, "Sample Survey Theory vs. General Statistical Theory: Estimation of the Population Mean," *International Statistical Review/Revue Internationale de Statistique*, vol. 40, no. 1, pp. 1–12, 1972.
- [18] M. Khan and J. Shabbir, "Some improved ratio, product, and regression estimators of finite population mean when using minimum and maximum values," *The Scientific World Journal*, vol. 2013, pp. 1–7, 2013.
- [19] D. N. Gujarati, *Basic Econometrics*, Tata McGraw-Hill Education, Gujarati, 2009.
- [20] P. Mukhopadhyay, *Small Area Estimation in Survey Sampling*, Alpha Science International, Oxford, UK, 1998.