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 Almigir Khalil, 5 Javid Shabbir, 3,6 and Wahidullah Mansoor 7

1 Department of Mathematics, College of Science and Humanities in Al-Khary, Prince Sattam Bin Abdulaziz University, Al-Khary 11942, Saudi Arabia
2 Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt
3 Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
4 Foundation Foundation University Medical College, Foundation University School of Health Sciences Islamabad, DHA-I Islamabad 44000, Pakistan
5 Department of Statistics, University of Peshawar, Peshawar, Pakistan
6 Department of Statistics, University of Wah, Wah Cantt, Pakistan
7 Department 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 practice 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 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 \(k = 12\) 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 simple 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 \(U = \{U_1, U_2, \ldots, U_N\}\) 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 \(y_{ij}\), \(x_{ij}\), \(z_{ij}\) where \(i = 1, 2, \ldots, n\) and \(j = 1, 2, \ldots, k\), are the values of \(i^{th}\) unit in the \(j^{th}\) selected sample for \(Y, X,\) and \(Z\) variables correspondingly. For \(y, x,\) and \(z\), the sample mean in systematic random sampling is

\[
\bar{Y}_{sys} = \frac{1}{n} \sum_{j=1}^{n} y_{ij},
\]

\[
\bar{X}_{sys} = \frac{1}{n} \sum_{j=1}^{n} x_{ij},
\]

\[
\bar{Z}_{sys} = \frac{1}{n} \sum_{j=1}^{n} z_{ij},
\]

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

\[
e_0 = \frac{\bar{Y}_{sys} - \bar{Y}}{\bar{Y}},
\]

\[
e_1 = \frac{\bar{X}_{sys} - \bar{X}}{\bar{X}},
\]

\[
e_2 = \frac{\bar{Z}_{sys} - \bar{Z}}{\bar{Z}},
\]

\[e_0 = \frac{\bar{Y}_{sys} - \bar{Y}}{\bar{Y}},
\]

\[e_1 = \frac{\bar{X}_{sys} - \bar{X}}{\bar{X}},
\]

\[e_2 = \frac{\bar{Z}_{sys} - \bar{Z}}{\bar{Z}},
\]

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

\[
E(e_0^2) = \lambda \rho_y^2 C_y^2,
\]

\[
E(e_1^2) = \lambda \rho_x^2 C_x^2,
\]

\[
E(e_2^2) = \lambda \rho_z^2 C_z^2,
\]

\[
E(e_0 e_1) = \lambda \rho_y \rho_x C_y C_x \sqrt{\rho_y^* \rho_x^*},
\]

\[
E(e_0 e_2) = \lambda \rho_y \rho_z C_y C_z \sqrt{\rho_y^* \rho_z^*},
\]

\[
E(e_1 e_2) = \lambda \rho_x \rho_z C_x C_z \sqrt{\rho_x^* \rho_z^*}.
\]

where

\[
\rho_y^* = 1 + (n - 1) \rho_y,
\]

\[
\rho_x^* = 1 + (n - 1) \rho_x,
\]

\[
\rho_z^* = 1 + (n - 1) \rho_z,
\]

\[
C_y = \frac{S_y}{\bar{Y}}
\]

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

\[
C_x = \frac{S_x}{\bar{X}}
\]

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

\[
C_z = \frac{S_z}{\bar{Z}}
\]

\[
S_z = \sqrt{\frac{1}{N - 1} \sum_{i=1}^{N} (z_i - \bar{Z}^2)}.
\]
The coefficients of $y$, $x$, and $z$ are $C_y$, $C_x$, and $C_z$, respectively, and

\[
\rho_y = \frac{E[(y_{ij} - \bar{y})(y_{ij} - \bar{y})]}{E(y_{ij} - \bar{y})^2},
\]

\[
\rho_x = \frac{E[(x_{ij} - \bar{x})(x_{ij} - \bar{x})]}{E(x_{ij} - \bar{x})^2},
\]

\[
\rho_z = \frac{E[(z_{ij} - \bar{z})(z_{ij} - \bar{z})]}{E(z_{ij} - \bar{z})^2},
\]

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

\[
\rho_{yx} = \frac{E(y_{ij} - \bar{y})(x_{ij} - \bar{x})}{\sqrt{E(y_{ij} - \bar{y})^2 E(x_{ij} - \bar{x})^2}},
\]

\[
\rho_{xz} = \frac{E(x_{ij} - \bar{x})(z_{ij} - \bar{z})}{\sqrt{E(x_{ij} - \bar{x})^2 E(z_{ij} - \bar{z})^2}},
\]

\[
\rho_{yz} = \frac{E(y_{ij} - \bar{y})(z_{ij} - \bar{z})}{\sqrt{E(y_{ij} - \bar{y})^2 E(z_{ij} - \bar{z})^2}},
\]

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:

\[
\tilde{y} = \begin{cases} 
\tilde{y} + c, & \text{if samples contain } y_{\min} \text{ but not } y_{\max} \\
\tilde{y} - c, & \text{if samples contain } y_{\max} \text{ but not } y_{\min} \\
\tilde{y}, & \text{for all other samples.}
\end{cases}
\]

The variance of $\tilde{y}$, is calculated as follows:

\[
\text{Var}(\tilde{y}) = \frac{2\lambda c}{N - 1} (y_{\max} - y_{\min} - nc),
\]

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

\[
c_{\text{opt}} = \frac{(y_{\max} - y_{\min})^2}{2n},
\]

The variance of $\tilde{y}$, is calculated as follows:

\[
\text{Var}(\tilde{y}) = V(\tilde{y}) - \lambda (y_{\max} - y_{\min})^2 \frac{2}{2(N - 1)},
\]

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

Under a systematic sampling approach, the usual ratio estimator is

\[
\tilde{y}_{\text{sys}} = \tilde{y} \frac{X}{X_{\text{sys}}},
\]

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

\[
\text{Bias}(\tilde{y}_{\text{sys}}) = \tilde{y} \lambda \left[ C_x \rho_x^* + C_z \rho_z^* - \rho_{yx} C_x C_z \rho_y^* \rho_x^* \right. \\
\left. - \rho_{xz} C_x C_z \sqrt{\rho_y^* \rho_z^*} + \rho_{xz} C_x C_z \sqrt{\rho_y^* \rho_z^*} \right],
\]

and

\[
\text{MSE}(\tilde{y}_{\text{sys}}) = \tilde{y} \lambda \left[ S_y^2 + R_1 S_y^2 \rho_x^* + R_2 S_y^2 \rho_z^* - 2 R_3 S_y \sqrt{\rho_y^* \rho_x^*} \\
+ 2 R_1 R_2 S_{xz} \sqrt{\rho_y^* \rho_z^*} - 2 R_3 S_y \sqrt{\rho_y^* \rho_z^*} \right].
\]

where, $R_1 = \frac{\tilde{y}}{X}$ and $R_2 = \frac{\tilde{y}}{X_{\text{sys}}}$. Under systematic sampling, the product estimator is as follows:

\[
\tilde{y}_{\text{sys}} = \tilde{y} \frac{X_{\text{sys}}}{X},
\]

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

\[
\text{Bias}(\tilde{y}_{\text{sys}}) = \tilde{y} \lambda \left[ \rho_{yx} C_x C_z \sqrt{\rho_y^* \rho_x^*} + \rho_{xz} C_x C_z \sqrt{\rho_y^* \rho_z^*} \\
+ \rho_{xz} C_x C_z \sqrt{\rho_y^* \rho_z^*} \right],
\]

and

\[
\text{MSE}(\tilde{y}_{\text{sys}}) = \tilde{y} \lambda \left[ S_y^2 + R_1 S_y^2 \rho_x^* + R_2 S_y^2 \rho_z^* - 2 R_3 S_y \sqrt{\rho_y^* \rho_x^*} \\
+ 2 R_1 R_2 S_{xz} \sqrt{\rho_y^* \rho_z^*} - 2 R_3 S_y \sqrt{\rho_y^* \rho_z^*} \right].
\]

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

\[
\tilde{y}_{\text{sys}} = \tilde{y} + b_1 \left( \bar{x} - \bar{x}_{\text{sys}} \right) + b_2 \left( \bar{z} - \bar{z}_{\text{sys}} \right).
\]

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 $\tilde{y}_{\text{sys}}$ is as follows:
\[
V\left(\hat{\theta}_{tr(sys)}\right) = \lambda S^2 \rho^*_y\left[1 - \left(\rho^2_{yx} + \rho^2_{yz} - 2\rho_{yx}\rho_{yz}\rho_{xz}\right)\right]. \tag{18}
\]

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.

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X}}{\hat{X}_{\text{sys}}}\right) \left(\frac{\hat{Z}}{\hat{Z}_{\text{sys}}}\right). \tag{19}
\]

or

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X} + c_1}{\hat{X}_{\text{sys}} + c_2}\right) \left(\frac{\hat{Z} + c_3}{\hat{Z}_{\text{sys}} + c_3}\right), \quad \text{if sample contains } y_{\min} \text{ and } x_{\min}, z_{\min}\]

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X} - c_1}{\hat{X}_{\text{sys}} - c_2}\right) \left(\frac{\hat{Z} - c_3}{\hat{Z}_{\text{sys}} - c_3}\right), \quad \text{if sample contains } y_{\max} \text{ and } x_{\max}, z_{\max}\]

for all the rest the 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

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X}_{\min}}{\hat{X}_{\text{sys}}}\right) \left(\frac{\hat{X}_{\max}}{\hat{X}_{\text{sys}}}\right). \tag{20}
\]

or

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X}_{\text{sys}} + c_1}{\hat{X}_{\text{sys}} + c_2}\right) \left(\frac{\hat{X}_{\text{sys}} + c_3}{\hat{X}_{\text{sys}} + c_3}\right), \quad \text{if sample contains } y_{\min} \text{ and } x_{\min}, z_{\min}\]

\[
\hat{\gamma}_{\text{suggested(sys)}} = \hat{\gamma}_{\text{sys}} \left(\frac{\hat{X}_{\text{sys}} - c_1}{\hat{X}_{\text{sys}} - c_2}\right) \left(\frac{\hat{X}_{\text{sys}} - c_3}{\hat{X}_{\text{sys}} - c_3}\right), \quad \text{if sample contains } y_{\max} \text{ and } x_{\max}, z_{\max}\]
The estimator of regression type is as follows:
\[
\hat{Y}_{\text{suggested}(\text{sys})} = \hat{Y}_\text{sys} + c_1 \left( \frac{\hat{Y}_\text{sys} + c_2}{\hat{Y}} \right) \left( \hat{Y}_\text{sys} + c_3 \right),
\]
for all other samples.

(23)

where \( \hat{Y}_\text{sys} = \hat{Y} + c_1, \hat{Y}_{c_2} = \hat{Y} - c_2, \hat{Y}_{c_3} = \hat{Y} + c_3 \). If the samples contain \( y_{\min} \) and \( x_{\max} \), and also \( c_1, c_2, c_3 \) 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.

\[
e_0 = \frac{\hat{Y}_\text{sys} - \hat{Y}_\text{sys}}{\hat{Y}} ,
\]

\[
e_1 = \frac{\hat{Y}_\text{sys} - \hat{Y}}{\hat{Y}} ,
\]

\[
e_2 = \frac{\hat{Y}_\text{sys} - \hat{Y}}{\hat{Y}} ,
\]

such that

\[
E(e_0) = E(e_1) = E(e_2) = 0
\]

\[
E(e_1^2) = \frac{\hat{X}^2}{\hat{X}^2} \left[ \frac{2\sigma_1^2}{N-1} (\Delta_y - nc_1) \right],
\]

\[
E(e_2^2) = \frac{\hat{X}^2}{\hat{X}^2} \left[ \frac{2\sigma_2^2}{N-1} (\Delta_z - nc_2) \right],
\]

\[
E(e_3^2) = \frac{\hat{X}^2}{\hat{X}^2} \left[ \frac{2\sigma_3^2}{N-1} (\Delta_z - nc_3) \right],
\]

\[
E(e_0 e_1) = \frac{\hat{X}^2}{\hat{X}^2} \left\{ S_{xy} \rho_{yz}^* - \frac{n}{N-1} [c_2 \Delta_y + c_1 \Delta_z - 2nc_1 c_2] \right\},
\]

\[
E(e_0 e_2) = \frac{\hat{X}^2}{\hat{X}^2} \left\{ S_{xz} \rho_{yz}^* - \frac{n}{N-1} [c_3 \Delta_y + c_1 \Delta_z - 2nc_2 c_3] \right\},
\]

\[
E(e_0 e_3) = \frac{\hat{X}^2}{\hat{X}^2} \left\{ S_{xz} \rho_{yz}^* - \frac{n}{N-1} [c_3 \Delta_y + c_1 \Delta_z - 2nc_2 c_3] \right\}.
\]

(24)

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

\[
\hat{Y}_{\text{suggested}(\text{sys})} = \hat{Y} (1 + e_0) (1 + e_1)^{-1} (1 + e_2)^{-1}.
\]

(25)

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

\[
\hat{Y}_{\text{suggested}(\text{sys})} - \hat{Y} = \hat{Y} (e_0 - e_1 - e_2 + e_1^2 + e_1 e_2 - e_0 e_2 - e_0 e_1).
\]

(26)

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

\[
\text{Bias} \left( \hat{Y}_{\text{suggested}(\text{sys})} \right) = \lambda \left[ \frac{R_1}{\hat{X}^2} \left\{ S_{xy}^2 \rho_{yz}^* - \frac{2nc_3}{N-1} (\Delta_z - nc_3) \right\} + R_2 \left\{ S_{xz}^2 \rho_{yz}^* - \frac{2nc_2}{N-1} (\Delta_z - nc_2) \right\} 
\right.
\]

\[
\left. - R_1 R_2 \left\{ S_{xz} \rho_{yz}^* - \frac{n}{N-1} [c_3 \Delta_y + c_1 \Delta_z - 2nc_2 c_3] \right\} \right] \right].
\]

(27)

(28)

(29)
We have after squaring (28) and taking expectations

\[
\text{MSE}\left( \widehat{Y}_{R\text{\ suggested}}(\text{sys}) \right) \equiv \left[ \lambda \left( S^2_{\beta Y} + R^2_{12\beta Y} + R^2_{22\beta Y} + 2 R_1 R_2 S_{YX} \sqrt{\beta^*_Y \beta^*_z} - 2 R_2 S_{YZ} \sqrt{\beta^*_Y \beta^*_z} \right) \right] - 2n\lambda \frac{(c_1 R_2 c_3 - c_1 c_2)}{N-1} \left\{ \Delta_y - R_1 \Delta_x - R_2 \Delta_z - n(c_1 - R_2 c_3 - R_1 c_2) \right\}
\]

(30)

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

\[
c_{1\text{opt}} = \frac{y_{\text{max}} - y_{\text{min}}}{2n},
\]

\[
c_{2\text{opt}} = \frac{x_{\text{max}} - x_{\text{min}}}{2n},
\]

\[
c_{3\text{opt}} = \frac{\sigma_{\text{max}} - \sigma_{\text{min}}}{2n},
\]

(31)

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

\[
\text{MSE}\left( \widehat{Y}_{R\text{\ suggested}}(\text{sys}) \right)_{\min} \equiv \text{MSE}\left( \widehat{Y}_{R\text{\ (sys)}} \right) = -\frac{\lambda}{2(N-1)} \left[ \Delta_y - R_1 \Delta_x - R_2 \Delta_z \right]^2,
\]

(32)

\[
\text{Bias}\left( \widehat{Y}_{R\text{\ suggested}}(\text{sys}) \right) \equiv \lambda \left[ \frac{R_1}{N-1} \left\{ S^2_{\beta Y} - \frac{2nc_3}{N-1} \Delta_x - nc_3 \right\} + R_2 \left\{ S^2_{\beta z} - \frac{2nc_3}{N-1} \Delta_z - nc_3 \right\} \right]
\]

\[
+ R_1 R_2 \left\{ S_{YZ} \sqrt{\beta^*_Y \beta^*_z} - \frac{n}{N-1} \left[ c_3 \Delta_y + c_2 \Delta_z - 2nc_3 c_3 \right] \right\}
\]

\[
+ R_1 \left\{ S_{YX} \sqrt{\beta^*_Y \beta^*_z} - \frac{n}{N-1} \left[ c_3 \Delta_y + c_2 \Delta_z - 2nc_3 c_3 \right] \right\}
\]

\[
+ R_2 \left\{ S_{XX} \sqrt{\beta^*_Y \beta^*_z} - \frac{n}{N-1} \left[ c_3 \Delta_y + c_2 \Delta_z - 2nc_3 c_3 \right] \right\}
\]

(34)

\[
\text{MSE}\left( \widehat{Y}_{P\text{\ suggested}}(\text{sys}) \right) \equiv \text{MSE}\left( \widehat{Y}_{P\text{\ suggested}}(\text{sys}) \right)_{\min} = -\frac{\lambda}{2(N-1)} \left[ \Delta_y + R_1 \Delta_x + R_2 \Delta_z \right]^2,
\]

(35)

where

\[
\text{MSE}\left( \widehat{Y}_{R\text{\ suggested}}(\text{sys}) \right) \equiv \text{MSE}\left( \widehat{Y}_{R\text{\ (sys)}} \right) = \lambda \left[ S^2_{\beta Y} + R^2_{12\beta Y} + R^2_{22\beta Y} - 2 R_1 R_2 S_{YX} \sqrt{\beta^*_Y \beta^*_z} \right] + 2 R_1 R_2 S_{XX} \sqrt{\beta^*_Y \beta^*_z}
\]

(33)

Similarly

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

\[
\text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{lr}} \leq \text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{sys}}
\]

\[
\lambda \frac{\Delta_y + \beta_1 \Delta_x + \beta_2 \Delta_z}{2(N-1)} \geq 0,
\]

\[
\text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{lr}} \leq \text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{sys}}
\]

\[
\lambda \frac{\Delta_y + \beta_1 \Delta_x + \beta_2 \Delta_z}{2(N-1)} \geq 0.
\]

Condition (ii).

By (16) and (33)

\[
\text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{lr}} \leq \text{MSE}(\hat{\beta}_{\text{suggested}})_{\text{sys}}
\]

\[
\lambda \frac{\Delta_y - \beta_1 \Delta_x - \beta_2 \Delta_z}{2(N-1)} \geq 0.
\]

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 index, all items (1967 = 100).

\( N = 12, n = 5, \)

\( \overline{y} = 132.7322, \overline{x} = 1781.125, \overline{z} = 174.075, \)

\( C_y = 0.5665, C_x = 0.3451, C_z = 0.2794, S_y = 75.19282, \)

\( S_x = 614.6662, S_z = 48.63655, \rho_{yx} = 0.9949, \)}
\[ \rho_{yz} = 0.9916, \quad \rho_{zx} = 0.9949, \quad \rho_y = 0.2867, \quad \rho_x = 0.2991, \]
\[ \rho_z = 0.3421, \quad S_{yx} = 0.08535, \quad S_{sx} = 0.10118, \quad S_{zx} = 0.0972, \]
\[ y_{\text{max}} = 265.086, \quad y_{\text{min}} = 39.866, \quad x_{\text{max}} = 2957.8, \]
\[ x_{\text{min}} = 992.7, \quad z_{\text{max}} = 272.41, \quad z_{\text{min}} = 116.3. \]

Population 2: (Source: [20])
\[ y = \text{Production output from 40 factories.} \]
\[ x = \text{The total number of employees.} \]
\[ z = \text{Fixed capital.} \]
\[ N = 40, \quad n = 10, \]
\[ \mathcal{F} = 5078.575, \quad \mathcal{F} = 230.325, \quad \mathcal{F} = 945.425, \quad C_y = 0.329525, \quad C_z = 0.84056, \quad C_x = 0.67555, \quad S_y = 1673.517, \]
\[ S_z = 193.602, \quad S_x = 638.6819, \quad \rho_{yx} = 0.8005, \quad \rho_{xz} = 0.83491, \quad \rho_{zx} = 0.98161, \quad \rho_y = 0.23, \quad \rho_x = 0.086, \]
\[ \rho_z = 0.16, \quad y_{\text{max}} = 8512, \quad y_{\text{min}} = 1451, \quad x_{\text{max}} = 662, \]
\[ x_{\text{min}} = 52, \quad z_{\text{max}} = 2355, \quad z_{\text{min}} = 107. \]

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