Hindawi Publishing Corporation
The Scientific World Journal
Volume 2013, Article ID 431868, 7 pages
http://dx.doi.org/10.1155/2013/431868

Research Article

Some Improved Ratio, Product, and Regression Estimators of Finite Population Mean When Using Minimum and Maximum Values

Manzoor Khan and Javid Shabbir

Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan

Correspondence should be addressed to Javid Shabbir; js@qau.edu.pk

Received 5 August 2013; Accepted 16 September 2013

Academic Editors: T. Robak and Y. Xia

Efficient estimation of finite population mean is carried out by using the auxiliary information meaningfully. In this paper we have suggested some modified ratio, product, and regression type estimators when using minimum and maximum values. Expressions for biases and mean squared errors of the suggested estimators have been derived up to the first order of approximation. The performances of the suggested estimators, relative to their usual counterparts, have been studied, and improved performance has been established. The improvement in efficiency by making use of maximum and minimum values has been verified numerically.

1. Introduction

Supplementary information in form of the auxiliary variable is rigorously used for the estimation of finite population mean for the study variable. Ratio and product estimators due to Cochran [1] and Murthy [2], respectively, are good examples when information on the auxiliary variable is incorporated for improved estimation of finite population mean of the study variable. When correlation between the study variable (y) and the auxiliary variable (x) is positive, ratio method of estimation is effective and when correlation is negative, product method of estimation is used. There are a lot of improvements and advancements in the construction of ratio, product, and regression estimators using the auxiliary information. For recent details, see Haq et al. [3], Haq and Shabbir [4], Yadav and Kadilar [5], Kadilar and Cingi [6], and Koyuncu and Kadilar [7] and the references cited therein.

The ratio method of estimation is at its best when the relationship between y and x is linear and the line of regression passes through the origin but as the line departs from origin, the efficiency of this method decreases. In practice, the condition that the line of regression passes through the origin is rarely satisfied and regression estimator is used for estimation of population mean.

Let $U = (U_1, U_2, \ldots, U_N)$ be a population of size N. Let (y_i, x_i) be the values of the study and the auxiliary variables, respectively, on the ith unit of a finite population.

Let us assume that a simple random sample of size n is drawn without replacement from U for estimating the population mean $\overline{Y} = \sum_{i=1}^{N} y_i/N$. It is further assumed that the population mean $\overline{X} = \sum_{i=1}^{N} x_i/N$ of the auxiliary variable x is known. The minimum say ($x_{\min}$) and maximum say ($x_{\max}$) values of the auxiliary variables are also assumed to be known.

The variance of mean per unit estimator $\overline{y} = \sum_{i=1}^{n} y_i/n$ is given by

$$V(\overline{y}) = \lambda S_y^2, \tag{1}$$

where $\lambda = ((1/n) - (1/N))$ and $S_y^2 = (1/(N-1)) \sum_{i=1}^{N} (y_i - \overline{Y})^2$.

Some time there exists unusually very large (say $y_{\max}$) and very small (say $y_{\min}$) units in the population. The mean per unit estimator is very sensitive to these unusual observations and as a result population mean will be either underestimated (in case the sample contains $y_{\min}$) or overestimated

(in case the sample contains $y_{\max}$). To overcome the situation Sarndal [8] suggested the following unbiased estimator:

$$\overline{y}_s = \begin{cases} \overline{y} + c & \text{if sample contains } y_{\min} \text{ but not } y_{\max}, \\ \overline{y} - c & \text{if sample contains } y_{\max} \text{ but not } y_{\min}, \\ \overline{y} & \text{for all other samples,} \end{cases} \tag{2}$$

where c is a constant.

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

$$V(\overline{y}_s) = \lambda S_y^2 - \frac{2\lambda nc}{N-1}(y_{\max} - y_{\min} - nc). \tag{3}$$

Further, $V(\overline{y}_s) < V(\overline{y})$ if $0 < c < (y_{\max} - y_{\min})/n$.

For, $c_{\text{opt}} = (y_{\max} - y_{\min})/2n$, variance of $\overline{y}_s$ is given by

$$V(\overline{y}_s)_{\text{opt}} = V(\overline{y}) - \frac{\lambda(y_{\max} - y_{\min})^2}{2(N-1)}, \tag{4}$$

which is always smaller than $V(\overline{y})$.

The usual ratio and product estimators of population mean $(\overline{Y})$ are given by

$$\overline{y}_R = \overline{y}\left(\frac{\overline{X}}{\overline{x}}\right), \qquad \overline{y}_P = \overline{y}\left(\frac{\overline{x}}{\overline{X}}\right), \tag{5}$$

where $\overline{y} = \sum_{i=1}^{n} y_i/n$ and $\overline{x} = \sum_{i=1}^{n} x_i/n$ are the sample means of variables y and x, respectively.

The expressions for biases $(B(\cdot))$, and mean square errors $(M(\cdot))$, of the conventional ratio and product estimators, are given by

$$B(\overline{y}_R) \cong \overline{Y}\lambda\left(C_x^2 - \rho_{yx}C_yC_x\right), \tag{6}$$

$$B(\overline{y}_P) = \overline{Y}\lambda\rho_{yx}C_yC_x, \tag{7}$$

$$M(\overline{y}_R) \cong \overline{Y}^2\lambda\left(C_y^2 + C_x^2 - 2\rho_{yx}C_yC_x\right), \tag{8}$$

$$M(\overline{y}_p) \cong \overline{Y}^2\lambda\left(C_y^2 + C_x^2 + 2\rho_{yx}C_yC_x\right), \tag{9}$$

where $C_y = S_y/\overline{Y}$ and $C_x = S_x/\overline{X}$ are the coefficients of variation of y and x, respectively, $\rho_{yx} = S_{yx}/S_yS_x$ is the correlation coefficient between y and x, $S_x^2 = \sum_{i=1}^{N}(x_i - \overline{X})^2/(N-1)$, and $S_{yx} = \sum_{i=1}^{N}(y_i - \overline{Y})(x_i - \overline{X})/(N-1)$ are the population variance and population covariance, respectively.

Usual regression estimator is given by

$$\overline{y}_{lr} = \overline{y} + b\left(\overline{X} - \overline{x}\right), \tag{10}$$

where b is the sample regression coefficient.

The variance of the estimator $\overline{y}_{lr}$ is given by

$$V(\overline{y}_{lr}) = \lambda S_y^2\left(1 - \rho_{yx}^2\right). \tag{11}$$

2. Proposed Estimators

Motivated by Sarndal [8], we extend this idea to estimators which make use of the auxiliary information for increased precision. It is well known that ratio and product estimators are used when y and x are positively and negatively correlated, respectively. We suggest estimator for each case separately as follows.

Case 1 (positive correlation between y and x). When y and x are positively correlated, then with selection of a larger value of x, a larger value of y is expected to be selected and when smaller value of x is selected, selection of a smaller value of y is expected. So we define the following estimators:

$$\widehat{\overline{Y}}_{RC} = \frac{\overline{y}_{c_{11}}}{\overline{x}_{c_{21}}}\overline{X} = \begin{cases} \dfrac{\overline{y} + c_1}{\overline{x} + c_2}\overline{X}, \\ \dfrac{\overline{y} - c_1}{\overline{x} - c_2}\overline{X}, \\ \dfrac{\overline{y}}{\overline{x}}\overline{X} \end{cases} \tag{12}$$

and similarly

$$\overline{y}_{lrC1} = \overline{y}_{c_{11}} + b\left(\overline{X} - \overline{x}_{c_{21}}\right), \tag{13}$$

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

Case 2 (negative correlation between y and x). When y and x are negatively correlated then with selection of a larger value of x, a smaller value of y is expected to be selected and when smaller value of x is selected, a larger value of y is expected to be selected. Keeping these points in view, the following estimators are therefore suggested:

$$\widehat{\overline{Y}}_{PC} = \overline{y}_{c_{12}}\frac{\overline{x}_{c_{22}}}{\overline{X}} = \begin{cases} \dfrac{(\overline{y} + c_1)(\overline{x} - c_2)}{\overline{X}}, \\ \dfrac{(\overline{y} - c_1)(\overline{x} + c_2)}{\overline{X}}, \\ \dfrac{\overline{y}}{\overline{x}}\overline{X} \end{cases} \tag{14}$$

and similarly

$$\overline{y}_{lrC2} = \overline{y}_{c_{11}} + b\left(\overline{X} - \overline{x}_{c_{22}}\right), \tag{15}$$

where $(\overline{y}_{c_{12}} = \overline{y} + c_1, \overline{x}_{c_{22}} = \overline{x} - c_2)$ if sample contains $y_{\min}$ and $x_{\max}$; $\overline{y}_{c_{12}} = (\overline{y} - c_1, \overline{x}_{c_{22}} = \overline{x} + c_2)$, if sample contains $y_{\max}$ and $x_{\min}$, and $(\overline{y}_{c_{12}} = \overline{y}, \overline{x}_{c_{22}} = \overline{x})$ for all other combinations of samples.

To find the bias and mean square error of these suggested estimators, we first prove two theorems which will be used in subsequent derivations.

Theorem 1. *If a sample of size n units is drawn from a population of size N units, then the covariance between $\overline{y}_{c_{12}}$ and $\overline{x}_{c_{12}}$, when they are positively correlated, is given by*

$$\operatorname{Cov}\left(\overline{y}_{c_{11}}, \overline{x}_{c_{21}}\right) = \lambda S_{yx} - \frac{n\lambda}{N-1} \times \left[c_1\left(x_{\max} - x_{\min}\right) + c_2\left(y_{\max} - y_{\min}\right) - 2nc_1c_2\right]. \tag{16}$$

Proof. Let us assume that n units have been drawn without replacement from a population of size N. Let S_n denote a sample space. We partition the whole sample space into three mutually exclusive and collectively exhaustive sets, that is, S_1, S_2, and S_3 such that $S_n = S_1 \cup S_2 \cup S_3$. Further S_1 is the set of all possible samples which contains $y_{\min}$ and $x_{\min}$, and S_2 consists of all samples which contains $y_{\max}$ and $x_{\max}$, and $S_3 = S - S_1 - S_2$. The number of sample points in S_1, S_2, and S_3 is given by $\binom{N-2}{n-1}$, $\binom{N-2}{n-1}$, and $\binom{N}{n} - 2\binom{N-2}{n-1}$, respectively.

By definition of covariance, we have

$$\begin{aligned}
&\operatorname{Cov}\left(\overline{y}_{c_{11}}, \overline{x}_{c_{21}}\right) \\
&= \binom{N}{n}^{-1}\left[\sum_{s\in S_1}\left(\overline{y} + c_1 - \overline{Y}\right)\left(\overline{x} + c_2 - \overline{X}\right) + \sum_{s\in S_2}\left(\overline{y} - c_1 - \overline{Y}\right)\left(\overline{x} - c_2 - \overline{X}\right) + \sum_{s\in S_3}\left(\overline{y} - \overline{Y}\right)\left(\overline{x} - \overline{X}\right)\right] \\
&= \binom{N}{n}^{-1}\left[\sum_{s\in S_1}\left(\overline{y} - \overline{Y}\right)\left(\overline{x} - \overline{X}\right) + \sum_{s\in S_2}\left(\overline{y} - \overline{Y}\right)\left(\overline{x} - \overline{X}\right) + \sum_{s\in S_3}\left(\overline{y} - \overline{Y}\right)\left(\overline{x} - \overline{X}\right) - c_1\left\{\sum_{s\in S_2}\left(\overline{x} - \overline{X}\right) - \sum_{s\in S_1}\left(\overline{x} - \overline{X}\right)\right\} - c_2\left\{\sum_{s\in S_2}\left(\overline{y} - \overline{Y}\right) - \sum_{s\in S_1}\left(\overline{y} - \overline{Y}\right)\right\} + c_1c_2\left\{\sum_{s\in S_1} + \sum_{s\in S_2}\right\}\right] \\
&= \binom{N}{n}^{-1}\left[\sum_{s\in S}\left(\overline{y} - \overline{Y}\right)\left(\overline{x} - \overline{X}\right)\right] - \binom{N}{n}^{-1} \times \left[c_1\left(\frac{x_{\max} - x_{\min}}{n}\right)\binom{N-2}{n-1} - c_2\left(\frac{y_{\max} - y_{\min}}{n}\right)\binom{N-2}{n-1} + 2c_1c_2\binom{N-2}{n-1}\right] \\
&= \operatorname{Cov}\left(\overline{y}, \overline{x}\right) - \frac{N-n}{N(N-1)} \times \left[c_1\left(x_{\max} - x_{\min}\right) + c_2\left(y_{\max} - y_{\min}\right) - 2nc_1c_2\right] \\
&= \lambda S_{yx} - \frac{n\lambda}{(N-1)} \times \left[c_1\left(x_{\max} - x_{\min}\right) + c_2\left(y_{\max} - y_{\min}\right) - 2nc_1c_2\right].
\end{aligned} \tag{17}$$

□

Theorem 2. *If a sample of size n units is drawn from a population of size N units, then the covariance between $\overline{y}_{c_{12}}$ and $\overline{x}_{c_{22}}$, when they are negatively correlated, is given by*

$$\operatorname{Cov}\left(\overline{y}_{c_{12}}, \overline{x}_{c_{22}}\right) = \lambda S_{yx} + \frac{n\lambda}{N-1} \times \left[c_1\left(x_{\max} - x_{\min}\right) + c_2\left(y_{\max} - y_{\min}\right) - 2nc_1c_2\right]. \tag{18}$$

The above Theorem 2 can be proved similarly as Theorem 1.

We define the following relative error terms.

Let $e_0 = (\overline{y}_{c_1} - \overline{Y})/\overline{Y}$ and $e_1 = (\overline{x}_{c_2} - \overline{X})/\overline{X}$, such that

$$E\left(e_0\right) = E\left(e_1\right) = 0, \tag{19}$$

$$E\left(e_0^2\right) = \frac{\lambda}{\overline{Y}^2}\left[S_y^2 - \frac{2nc_1}{N-1}\left(y_{\max} - y_{\min} - nc_1\right)\right], \tag{20}$$

$$E\left(e_1^2\right) = \frac{\lambda}{\overline{X}^2}\left[S_x^2 - \frac{2nc_2}{N-1}\left(x_{\max} - x_{\min} - nc_2\right)\right], \tag{21}$$

$$E\left(e_0e_1\right) = \frac{\lambda}{\overline{Y}\,\overline{X}}\left[S_{yx} - \frac{n}{N-1} \times \left\{c_2\left(y_{\max} - y_{\min}\right) + c_1\left(x_{\max} - x_{\min}\right) - 2nc_1c_2\right\}\right]. \tag{22}$$

Expressing $\hat{\bar{Y}}_{RC}$ in terms of e's, we have

$$\hat{\bar{Y}}_{RC} = \bar{Y}(1+e_0)(1+e_1)^{-1}. \tag{23}$$

Expanding and rearranging right-hand side of (23), to first degree of approximation, we have

$$\left(\hat{\bar{Y}}_{RC} - \bar{Y}\right) \cong \bar{Y}\left(e_0 + e_1 - e_0 e_1 + e_1^2\right). \tag{24}$$

Using (24), the bias of $\hat{\bar{Y}}_{RC}$ is given by

$$\begin{aligned} B\left(\hat{\bar{Y}}_{RC}\right) \cong \frac{\lambda}{\bar{X}} \Bigg[& R\left(S_x^2 - \frac{2nc_2}{N-1}\left(x_{\max} - x_{\min} - nc_2\right)\right) \\ & - \Bigg\{ S_{yx} - \frac{n}{N-1} \\ & \times \left(c_2\left(y_{\max} - y_{\min}\right) + c_1\left(x_{\max} - x_{\min}\right)\right. \\ & \left. - 2nc_1c_2\right) \Bigg\} \Bigg], \end{aligned} \tag{25}$$

where $R = \bar{Y}/\bar{X}$.

Using (24), the mean square error of $\hat{\bar{Y}}_{RC}$, to the first degree of approximation, is given by

$$\begin{aligned} M\left(\hat{\bar{Y}}_{RC}\right) \cong \lambda \Bigg[& S_y^2 - \frac{2nc_1}{N-1}\left(y_{\max} - y_{\min} - nc_1\right) \\ & + R^2 \left\{ S_x^2 - \frac{2nc_2}{N-1}\left(x_{\max} - x_{\min} - nc_2\right) \right\} \\ & - 2R \Bigg\{ S_{yx} - \frac{n}{N-1} \\ & \times \left(c_2\left(y_{\max} - y_{\min}\right) + c_1\left(x_{\max} - x_{\min}\right)\right. \\ & \left. - 2nc_1c_2\right) \Bigg\} \Bigg] \end{aligned} \tag{26}$$

or

$$\begin{aligned} M\left(\hat{\bar{Y}}_{RC}\right) \cong M\left(\bar{y}_R\right) - \frac{2\lambda n}{N-1} \\ \times \left[\left(c_1 - Rc_2\right)\left\{\left(y_{\max} - y_{\min}\right) - R\left(x_{\max} - x_{\min}\right)\right.\right. \\ \left.\left. - n\left(c_1 - Rc_2\right)\right\}\right]. \end{aligned} \tag{27}$$

To find optimum values of c_1 and c_2, we differentiate (27) with respect to c_1 and c_2 as

$$\begin{aligned} \frac{\partial M\left(\hat{\bar{Y}}_{RC}\right)}{\partial c_1} = 0 \Longrightarrow & \left(y_{\max} - y_{\min}\right) - R\left(x_{\max} - x_{\min}\right) \\ & - 2n\left(c_1 - Rc_2\right) = 0, \\ \frac{\partial M\left(\hat{\bar{Y}}_{RC}\right)}{\partial c_2} = 0 \Longrightarrow & \left(y_{\max} - y_{\min}\right) - R\left(x_{\max} - x_{\min}\right) \\ & - 2n\left(c_1 - Rc_2\right) = 0 \end{aligned} \tag{28}$$

Here we have one equation with two unknowns so unique, solution is not possible, so we let $c_2 = (x_{\max} - x_{\min})/2n$, and then $c_1 = (y_{\max} - y_{\min})/2n$.

For optimum values of c_1 and c_2, the optimum mean square error of $\hat{\bar{Y}}_{RC}$ is given by

$$\begin{aligned} M\left(\hat{\bar{Y}}_{RC}\right)_{\text{opt}} \cong M\left(\bar{y}_R\right) - \frac{\lambda}{2(N-1)} \\ \times \left[\left(y_{\max} - y_{\min}\right) - R\left(x_{\max} - x_{\min}\right)\right]^2. \end{aligned} \tag{29}$$

Similarly the bias and mean square error or optimum mean square error of $\hat{\bar{Y}}_{PC}$ are, respectively, given by

$$\begin{aligned} B\left(\hat{\bar{Y}}_{PC}\right) \cong \frac{\lambda}{\bar{X}} \Bigg[& S_{yx} - \frac{n}{N-1} \\ & \times \left\{c_2\left(y_{\max} - y_{\min}\right) + c_1\left(x_{\max} - x_{\min}\right)\right. \\ & \left. - 2nc_1c_2\right\} \Bigg], \end{aligned} \tag{30}$$

$$\begin{aligned} M\left(\hat{\bar{Y}}_{PC}\right) \cong M\left(\bar{y}_P\right) - \frac{2\lambda n}{N-1} \\ \times \left[\left(c_1 + Rc_2\right)\left\{\left(y_{\max} - y_{\min}\right) + R\left(x_{\max} - x_{\min}\right)\right.\right. \\ \left.\left. - n\left(c_1 + Rc_2\right)\right\}\right]. \end{aligned} \tag{31}$$

For optimum values of c_1 and c_2, the optimum mean square error of $\hat{\bar{Y}}_{PC}$ is given by

$$\begin{aligned} M\left(\hat{\bar{Y}}_{PC}\right)_{\text{opt}} \cong M\left(\bar{y}_P\right) - \frac{\lambda}{2(N-1)} \\ \times \left[\left(y_{\max} - y_{\min}\right) - R\left(x_{\max} - x_{\min}\right)\right]^2. \end{aligned} \tag{32}$$

The variance of regression estimator $\bar{y}_{lrC1}$ in case of positive correlation is given by

$$\begin{aligned} V\left(\bar{y}_{lrC1}\right) = V\left(\bar{y}_{lr}\right) - \frac{2\lambda n}{N-1} \\ \times \left[\left(c_1 - \beta c_2\right)\left\{\left(y_{\max} - y_{\min}\right) - \beta\left(x_{\max} - x_{\min}\right)\right.\right. \\ \left.\left. - 2n\left(c_1 - \beta c_2\right)\right\}\right], \end{aligned} \tag{33}$$

where $\beta = \rho_{yx}(S_y/S_x)$ is the population regression coefficient of y on x.

For $c_2 = (x_{\max} - x_{\min})/2n$ and $c_1 = (y_{\max} - y_{\min})/2n$, optimum variance of $\bar{y}_{lrC1}$ is given by

$$\begin{aligned} V\left(\bar{y}_{lrC1}\right)_{\text{opt}} = V\left(\bar{y}_{lr}\right) - \frac{\lambda}{2(N-1)} \\ \times \left[\left(y_{\max} - y_{\min}\right) - \beta\left(x_{\max} - x_{\min}\right)\right]^2. \end{aligned} \tag{34}$$

For negative correlation, variance of the regression estimator $\overline{y}_{lrC2}$ is given by

$$V(\overline{y}_{lrC2}) = V(\overline{y}_{lr}) - \frac{2\lambda n}{N-1} \times [(c_1 + \beta c_2)\{(y_{\max} - y_{\min}) + \beta(x_{\max} - x_{\min}) - 2n(c_1 + \beta c_2)\}]. \tag{35}$$

For $c_2 = (x_{\max} - x_{\min})/2n$ and $c_1 = (y_{\max} - y_{\min})/2n$, optimum variance of $\overline{y}_{lrC2}$ is given by

$$V(\overline{y}_{lrC2})_{\text{opt}} = V(\overline{y}_{lr}) - \frac{\lambda}{2(N-1)} \times [(y_{\max} - y_{\min}) + \beta(x_{\max} - x_{\min})]^2. \tag{36}$$

So in general we can write $V(\overline{y}_{lrC})_{\text{opt}}$ as

$$V(\overline{y}_{lrC})_{\text{opt}} = V(\overline{y}_{lr}) - \frac{\lambda}{2(N-1)} \times [(y_{\max} - y_{\min}) - |\beta|(x_{\max} - x_{\min})]^2. \tag{37}$$

3. Comparison

The conditions under which the suggested estimators $\hat{\overline{Y}}_{RC}$, $\hat{\overline{Y}}_{PC}$, and $\overline{y}_{lrC}$ perform better than the usual mean per unit estimator and their usual counterpart is given below.

(a) Comparison of Proposed Ratio Type Estimator. A proposed estimator $\hat{\overline{Y}}_{RC}$ will perform better than

(i) mean per unit estimator (by (1) and (27)) if $V(\overline{y}) - M(\hat{\overline{Y}}_{RC}) > 0$ or if

$$\rho_{yx} > \frac{RS_x}{2S_y} - \frac{n}{RS_yS_x(N-1)} \times [(c_1 - Rc_2)\{(y_{\max} - y_{\min}) - R(x_{\max} - x_{\min}) - 2n(c_1 - Rc_2)\}]; \tag{38}$$

(ii) usual ratio estimator (by (8) and (27)) if $M(\overline{y}_R) - M(\hat{\overline{Y}}_{RC}) > 0$ or if

$$\min\left[Rc_2, Rc_2 - \left\{\frac{R(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right] < c_1 < \max\left[Rc_2, Rc_2 - \left\{\frac{R(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right]. \tag{39}$$

(b) Comparison of Proposed Product Type Estimator. A proposed product type estimator will perform better than

(iii) mean per unit estimator if $V(\overline{y}) - M(\hat{\overline{Y}}_{PC}) > 0$ (by (1) and (31)) or if

$$\rho_{yx} < -\frac{RS_x}{2S_y} + \frac{n}{RS_yS_x(N-1)} \times [(c_1 + Rc_2)\{(y_{\max} - y_{\min}) + R(x_{\max} - x_{\min}) - 2n(c_1 + Rc_2)\}]; \tag{40}$$

(iv) usual product estimator if $M(\overline{y}_P) - M(\hat{\overline{Y}}_{PC}) > 0$ (by (9) and (31)) or if

$$\min\left[-Rc_2, -Rc_2 + \left(\frac{R(x_{\max} - x_{\min}) + (y_{\max} - y_{\min})}{n}\right)\right] < c_1 < \max\left[-Rc_2, -Rc_2 + \left(\frac{R(x_{\max} - x_{\min}) + (y_{\max} - y_{\min})}{n}\right)\right]. \tag{41}$$

(c) Comparison of Proposed Regression Type Estimator. A proposed regression type estimator (positive correlation) will perform better than

(v) mean per unit estimator if $V(\overline{y}) - V(\overline{y}_{lrC1}) > 0$ (by (1) and (33)) or if

$$\rho_{yx}^2 > -\frac{2n}{S_y^2(N-1)} \times [(c_1 - \beta c_2)\{(y_{\max} - y_{\min}) - \beta(x_{\max} - x_{\min}) - 2n(c_1 - \beta c_2)\}]; \tag{42}$$

(vi) usual regression estimator if $V(\overline{y}_{lr}) - V(\overline{y}_{lrC1}) > 0$ (by (11) and (33)) or if

$$\min\left[\beta c_2, \beta c_2 - \left\{\frac{\beta(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right] < c_1 < \max\left[\beta c_2, \beta c_2 - \left\{\frac{\beta(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right]. \tag{43}$$

A proposed regression type estimator (negative correlation) will perform better than

(vii) mean per unit estimator if $V(\overline{y}) - V(\overline{y}_{lrC2}) > 0$ (by (1) and (35)) or if

$$\rho_{yx}^2 > -\frac{2n}{S_y^2(N-1)} \times [(c_1 + \beta c_2)\{(y_{\max} - y_{\min}) + \beta(x_{\max} - x_{\min}) - 2n(c_1 - \beta c_2)\}]; \tag{44}$$

Table 1: Numerical values of conditions (38)–(45).

Conditions	Population 1	Population 2	Population 3	Population 4
(38)	$0.99 > 0.592$	$-0.667 > 0.2529^*$	$0.980 > 0.478$	$0.981 > 0.525$
(39)	$0.214 < 0.287 < 0.360$	$0.442 < 1.125 < 1.807$	$22.23 < 26.16 < 30.094$	$22.92 < 24.5 < 26.07$
(40)	$0.99 < -0.592^*$	$-0.667 < -0.253$	$0.980 < -0.4783^*$	$0.981 < -0.525^*$
(41)	$-0.360 < 0.287 < 0.93$	$-0.442 < 1.125 < 2.692$	$-22.238 < 26.16 < 74.57$	$-26.07 < 24.5 < 75.07$
(42)	$0.980 > 0.000$	$0.459 > 0.000$	$0.9605 > 0.000$	$0.963 > 0.000$
(43)	$0.273 < 0.287 < 0.301$	$-0.773 < 1.125 < -0.592^*$	$22.78 < 26.16 < 29.549$	$24.36 < 24.5 < 24.63$
(44)	$0.980 > -1.913$	$0.459 > 0.272$	$0.9605 > -0.862$	$0.963 > -1.88$
(45)	$-0.30 < 0.287 < -0.15^*$	$0.592 < 1.125 < 4.026$	$-30.63 < 26.16 < -22.78^*$	$-24.36 < 24.5 < -21.2^*$

Note: *indicates that the condition is not satisfied.

Table 2: PRE of different estimators with respect to $\overline{y}$.

Estimator	Population 1	Population 2	Population 3	Population 4
$\overline{y}_R$	1742.030	51.514	2501.29	2442.965
$\overline{y}_P$	21.053	175.210	26.383	23.998
$\overline{y}_{lr}$	5115.818	184.699	2536.131	2763.749
$\widehat{\overline{Y}}_{RC}$	2319.293	54.325	2940.086	2502.692
$\widehat{\overline{Y}}_{PC}$	21.117	212.639	26.424	24.004
$\overline{y}_{lrC}$	5249.673	208.254	2856.83	2764.336

(viii) usual regression estimator if $V(\overline{y}_{lr}) - V(\overline{y}_{lrC1}) > 0$ (by (11) and (33)) or if

$$\min\left[-\beta c_2, -\beta c_2 + \left\{\frac{\beta(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right] < c_1 < \max\left[-\beta c_2, -\beta c_2 + \left\{\frac{\beta(x_{\max} - x_{\min}) - (y_{\max} - y_{\min})}{n}\right\}\right]. \tag{45}$$

(d) Comparison of Suggested Estimators for Optimum Values of c_1 and c_2 with Usual Estimators. For optimum values of c_1 and c_2, the proposed estimator will always perform better than usual mean per unit estimator and their usual counterparts (ratio, product and regression estimators).

4. Empirical Study

We consider the following datasets for numerical comparison.

Population 1 (Singh and Mangat [9, page 193]). Let y be the milk yield in kg after new food and let x be the yield in kg before new yield. $N = 27$, $n = 12$, $\overline{X} = 10.41111$, $\overline{Y} = 11.25185$, $y_{\max} = 14.8$, $y_{\min} = 7.9$, $x_{\max} = 14.5$, $x_{\min} = 6.5$, $S_y^2 = 4.103$, $S_x^2 = 4.931$, $S_{xy} = 4.454$, and $\rho_{yx} = 0.990$.

Population 2 (Singh and Mangat [9, page 195]). Let y be the weekly time (hours) spent in nonacademic activities and let x be the overall grade point average (4.0 bases). $N = 36$, $n = 12$, $\overline{X} = 2.798333$, $\overline{Y} = 14.77778$, $y_{\max} = 33$, $y_{\min} = 6$, $x_{\max} = 3.82$, $x_{\min} = 1.81$, $S_y^2 = 38.178$, $S_x^2 = 0.3504$, $S_{xy} = -2.477$, and $\rho_{yx} = -0.6772$.

Population 3 (Murthy [10, page 399]). Let y be the area under wheat crop in 1964 and let x be the area under wheat crop in 1963. $N = 34$, $n = 12$, $\overline{X} = 208.882$, $\overline{Y} = 199.441$, $y_{\max} = 634$, $y_{\min} = 6$, $x_{\max} = 564$, $x_{\min} = 5$, $S_y^2 = 22564.56$, $S_x^2 = 22652.05$, $S_{xy} = 22158.05$, and $\rho_{yx} = 0.980$.

Population 4 (Cochran [11, page 152]). Let y be population size in 1930 (in 1000) and x be the population size in 1920 (in 1000). $N = 49$, $n = 12$, $\overline{X} = 103.1429$, $\overline{Y} = 127.7959$, $y_{\max} = 634$, $y_{\min} = 46$, $x_{\max} = 507$, $x_{\min} = 2$, $S_y^2 = 15158.83$, $S_x^2 = 10900.42$, $S_{xy} = 12619.78$ and $\rho_{yx} = 0.98$.

The conditional values and results are given in Tables 1 and 2, respectively.

For percentage relative efficiency (PRE), we use the following expression:

$$\mathrm{PRE}(\overline{y}_i, \overline{y}) = \frac{V(\overline{y})}{V(\overline{y}_i) \text{ or } M(\overline{y}_i)} \times 100 \quad \text{for} \quad i = R, P, RC, PC, lr, lrC. \tag{46}$$

5. Conclusion

From Table 2, it is observed that the ratio estimator $\widehat{\overline{Y}}_{RC}$ is performing better than $\overline{y}_R$ in Populations 1, 3, and 4 because

of positive correlation. The product estimator $\widehat{\overline{Y}}_{PC}$ is better than $\overline{y}_P$ just in Population 2 because of negative correlation. The regression estimator $\overline{y}_{lrC}$ outperforms than all other considered estimators and is preferable.

Acknowledgments

The authors are thankful to the learned referees for their valuable suggestions and helpful comments in revising the manuscript.

References

[1] W. G. Cochran, "The estimation of the yields of cereal experiments by sampling for the ratio to total produce," *The Journal of Agriculture Science*, vol. 30, pp. 262–275, 1940.

[2] M. N. Murthy, "Product method of estimation," *Sankhya A*, vol. 16, pp. 69–74, 1964.

[3] A. Haq, J. Shabbir, and S. Gupta, "Improved exponential ratio type estimators in stratified sampling," *Pakistan Journal of Statistics*, vol. 29, no. 1, pp. 13–31, 2013.

[4] A. Haq and J. Shabbir, "Improved family of ratio estimators in simple and stratified random sampling," *Communication in Statistics: Theory and Methods*, vol. 42, no. 5, pp. 782–799, 2013.

[5] S. K. Yadav and C. Kadilar, "Improved class of ratio and product estimators," *Applied Mathematics and Computation*, vol. 219, no. 22, pp. 10726–10731, 2013.

[6] C. Kadilar and H. Cingi, "Improvement in estimating the population mean in simple random sampling," *Applied Mathematics Letters*, vol. 19, no. 1, pp. 75–79, 2006.

[7] N. Koyuncu and C. Kadilar, "Ratio and product estimators in stratified random sampling," *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2552–2558, 2009.

[8] C. E. Sarndal, "Sample survey theory vs general statistical theory: estimation of the population mean," *International Statistical Institute*, vol. 40, pp. 1–12, 1972.

[9] R. Singh and N. S. Mangat, *Elements of Survey Sampling*, Kluwer Academic Publishers, London, UK, 1996.

[10] M. N. Murthy, *Sampling Theory and Methods*, Statistical Publishing Society, Calcutta, India, 1967.

[11] W. G. Cochran, *Sampling Techniques*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1977.