A new improved difference-cum-exponential ratio type estimator in systematic sampling using two auxiliary variables

J Shabbir1*, S Masood2 and S Gupta3

1 Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
2 Department of Mathematics and Statistics, PMAS-Arid Agriculture University Rawalpindi, Pakistan.
3 Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, USA.

Submitted: 22 May 2018; Revised: 23 January 2019; Accepted: 27 September 2019

Abstract: A difference-cum-exponential ratio type estimator was proposed for estimating the finite population mean using two auxiliary variables under systematic sampling. Expressions for the biases and mean square errors (MSEs) were derived up to first order of approximation. It was observed that the proposed estimator is more efficient than the usual sample mean estimator, traditional ratio estimator, exponential-ratio estimator and many other recently proposed difference type estimators in terms of MSEs. Four datasets were used for efficiency comparisons.

Keywords: Bias, efficiency, MSE, systematic sampling.

INTRODUCTION

The use of auxiliary information in sample survey in an appropriate way may increase the precision of estimators by taking advantages of correlation between the study variable and the auxiliary variable. The ratio, product, exponential and regression estimators have frequently been used by many researchers in different forms either at the estimation stage or at designing stage or at both stages. In daily life information on single as well as two auxiliary variables are commonly used to enhance the precision of estimators. For example: (i) let \( y \) be the yield of a particular crop based on the area of a crop \( x \) and the amount of water utilisation \( z \); (ii) let \( y \) be the electricity consumed in the households (HHDs) based on income of the HHDs \( x \) and size of the HHDs \( z \) and (iii) let \( y \) be the average grades of the undergraduate students based on the number of study hours used by students \( x \) and intelligence coefficient (IQ) level of the students \( z \). In all these examples, we are interested to use the auxiliary information in parallel to the study variable. Madow and Madow (1944), Cochran (1946) and Gautschi (1957) were among the earlier contributors in the area of systematic sampling. Other authors who have contributed in this area include: Swain (1964), Kushwaha and Singh (1989), Banarasi et al. (1993), Singh and Singh (1998), Singh et al. (2011), Singh and Jatwa (2012), Tailor et al. (2013), Khan and Singh (2015), Khan (2016), Pal and Singh (2017), Kocyigit and Cingi (2017), Riaz et al. (2017), Tailor and Mishra (2018), Qureshi et al. (2018) and Mishra et al. (2018). The main objective of this study was to construct a new estimator by taking advantage of two auxiliary variables to improve the efficiency of the estimator in systematic sampling.

METHODOLOGY

Consider a finite population \( U = \{U_1, U_2, \ldots, U_n, \ldots, U_N\} \) of \( N \) ordered identifiable units. A sample of size \( n \) units is selected by first selecting an observation from the first \( k \) units and then selecting every \( k \)th unit thereafter. Thus, we can obtain \( k \) possible samples each of size \( n \) such that \( nk = N \). Let \( y_{ij} \) and \( (x_{ij}, z_{ij}) \) \((i = 1, 2, 3, \ldots, k; j = 1, 2, 3, \ldots, n)\) be the observed values of the study variable \( y \) and the auxiliary variables \( x \) and \( z \), respectively for the \( j \)th unit in the \( i \)th possible
sample. Let \( \bar{Y}^{(0)} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n_i} y_{ij}, \bar{X}^{(0)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} x_{ij} \) and 
\( \bar{Z}^{(0)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} z_{ij} \) be the systematic sample means corresponding to population means
\( \bar{Y} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n_i} y_{ij}, \bar{X} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} x_{ij} \) and 
\( Z = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} z_{ij}, \) respectively. Let \( \rho^*_1 = 1 + (n-1) \rho_{yy}, \rho^*_2 = 1 + (n-1) \rho_{xx} \) and 
\( \rho^*_3 = 1 + (n-1) \rho_{zz} \), where \( \rho_{yy} = \frac{E(y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})}{E(y_{ij} - \bar{Y})^2} \)
\( \rho_{xx} = \frac{E(x_{ij} - \bar{X})(x_{ij}' - \bar{X})}{E(x_{ij} - \bar{X})^2} \) are
the intra-class correlation coefficients between pairs of units within the same systematic sample for \( y, x \) and 
\( z, \) respectively. Let \( \rho_{yx}, \rho_{xy} \) and \( \rho_{xz} \) be the usual population correlation coefficients between \( y, x \) and 
\( (y, z) \) and \( (x, z), \) respectively. Let \( C_y = \frac{S_y}{\bar{Y}}, C_x = \frac{S_x}{\bar{X}} \) and 
\( C_z = \frac{S_z}{\bar{Z}} \) be the coefficients of variation of \( y, x \) and \( z, \) respectively, where
\[
S_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left( \sum_{j=1}^{n_i} (y_{ij} - \bar{Y})^2 \right)}, \quad S_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left( \sum_{j=1}^{n_i} (x_{ij} - \bar{X})^2 \right)}
\]
and \( S_z = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} \left( \sum_{j=1}^{n_i} (z_{ij} - \bar{Z})^2 \right)} \) be the population standard deviations for \( y, x \) and \( z, \) respectively.

We also define the following error terms.

Let \( \zeta_i^{(0)} = \frac{Y_i^{(0)} - \bar{Y}}{\bar{Y}}, \quad \zeta_i = \frac{Y_i - \bar{Y}}{\bar{Y}}, \quad \zeta_i^{(0)} = \frac{Z^{(0)} - \bar{Z}}{\bar{Z}}, \) such that
\( E(\zeta_i) = 0 \) for \( i = 0, 1, 2. \) Also \( E(\zeta_i^{(0)}) = \Phi \rho_i^{(0)} C, \quad E(\zeta_i^0) = \Phi \rho^* C, \quad E(\zeta_i^2) = \Phi C, \quad E(\zeta_i^0) = \Phi C, \) \( \sqrt{\rho_i}, \rho_{xy}, \) \( \rho_{xz}, \) and \( \sqrt{\rho_i}, \rho_{yy}, \) 
respectively, where \( C_y = \rho_{yy} C_y, C_x = \rho_{xy} C_x, C_z = \rho_{xz} C_z, \) and 
\( \Phi = \frac{N-1}{nn^2}. \)

Now we discuss some of the existing estimators.

The usual sample mean estimator is given by
\( \hat{Y}_0 = \bar{Y}^{(0)} \) \( \quad \ldots (1) \)
The variance of \( \hat{Y}_0 \) is given by
\( \text{Var}(\hat{Y}_0) = \Phi \bar{Y} \rho_y^2 C_y \) \( \quad \ldots (2) \)

**Single auxiliary variable**

(i) The traditional ratio estimator is given by
\( \hat{Y}_R = \bar{Y}^{(0)} \left( \frac{\bar{X}}{\bar{Y}^{(0)}} \right) \) \( \quad \ldots (3) \)
The bias and MSE of \( \hat{Y}_R \) to first order of approximation are respectively given by
\( B(\hat{Y}_R) \equiv \Phi \bar{Y} \left( \rho_y^2 C_y^2 - C_y \sqrt{\rho_y^2} \rho_y \right) \) \( \quad \ldots (4) \)
and
\( \text{MSE}(\hat{Y}_R) \equiv \Phi \bar{Y} \left( \rho_y^2 C_y^2 + 2 C_y \sqrt{\rho_y^2} \rho_y \right) \) \( \quad \ldots (5) \)
(ii) The usual exponential ratio type estimator is given by
\( \hat{Y}_RE = \bar{Y}^{(0)} \exp \left( \frac{\bar{X} - \bar{Y}^{(0)}}{\bar{X} + \bar{Y}^{(0)}} \right) \) \( \quad \ldots (6) \)
The bias and MSE of \( \hat{Y}_RE \) to first order of approximation are respectively given by
\( B(\hat{Y}_RE) \equiv \Phi \bar{Y} \left( \frac{3}{8} \rho_y^2 C_y^2 - \frac{1}{2} C_y \sqrt{\rho_y^2} \rho_y \right) \) \( \quad \ldots (7) \)
and
\( \text{MSE}(\hat{Y}_RE) \equiv \Phi \bar{Y} \left( \rho_y^2 C_y^2 + \frac{1}{4} \rho_y^2 C_y^2 - C_y \sqrt{\rho_y^2} \rho_y \right) \) \( \quad \ldots (8) \)
(iii) The usual difference estimator is given by
\( \hat{Y}_D = \bar{Y}^{(0)} + d \left( \bar{X} - \bar{Y}^{(0)} \right), \) \( \quad \ldots (9) \)
where \( d \) is the constant.
The minimum variance of \( \hat{Y}_D \) at optimum value of \( d \)
i.e. \( d_{\text{opt}} = \frac{\bar{Y} \rho_y C_y \sqrt{\rho_y^2}}{\bar{X} C_y \rho_y} \), is given by
\( \text{Var}(\hat{Y}_{D,\text{opt}}) \equiv \Phi \bar{Y} \rho_y^2 C_y \left( 1 - \frac{1}{4} \rho_y \right) = \text{MSE}(\hat{Y}_{D,\text{opt}}) \) \( \quad \ldots (10) \)
(iv) Kocyigit and Cingi (2017) suggested a new class of
unbiased linear estimator which is given by

\[ \hat{Y}_{RC} = \sum_{i=1}^{2} \left( q_1 + q_2 \left( \frac{aX_{sys} + b}{a + b} \right) + q_3 \left( \frac{aX_{sys} + b}{a + b} \right)^3 \right), \]

where \((\alpha, \delta) \in \mathbb{R}\) and \(q(i = 1, 2, 3)\) are constants and \(a\) and \(b\) are the functions of known population parameters, which may be the population mean, population coefficient of variation, population coefficient of skewness and population coefficient of kurtosis of the auxiliary variable \(x\).

The minimum \(MSE\) of \(\hat{Y}_{RC}\) at optimum value of \(A_4 = (q_2 \alpha + q_3 \delta)\) i.e. \(A_{4(\text{opt})} = -\frac{\rho_{xy}C^2}{gC_y \rho_x}\), where \(g = \frac{aX}{a + b}\), is given by

\[ MSE(\hat{Y}_{RC})_{\text{min}} = \Phi \left( \frac{\sum \phi X - \sum \phi X_{sys}}{\sum \phi X + \sum \phi X_{sys}} \right)^2, \]

which is equal to the variance of the linear regression estimator.

(v) Riaz et al. (2017) suggested the following class of estimators in systematic sampling:

\[ \hat{Y}_{RDS} = \left\{ t_1 \sum \phi_{sys} + t_2 \left( \bar{X} - \sum \phi_{sys} \right) \right\} \exp \left( \frac{\sum \phi X - \sum \phi X_{sys}}{\sum \phi X + \sum \phi X_{sys}} \right), \]

where \(t(i = 1, 2)\) are constants and \(\gamma\) is the scalar quantity.

For \(\gamma = 1\), the above estimator becomes:

\[ \hat{Y}_{RDS} = \left\{ t_1 \sum \phi_{sys} + t_2 \left( \bar{X} - \sum \phi_{sys} \right) \right\} \exp \left( \frac{\sum \phi X - \sum \phi X_{sys}}{\sum \phi X + \sum \phi X_{sys}} \right), \]

The bias of \(\hat{Y}_{RDS}\) to first order of approximation is given by

\[ B(\hat{Y}_{RDS}) \equiv (t_1 - 1) \Phi + t_2 \Phi \left\{ \sum \phi X - \sum \phi X_{sys} \right\} \]

\[ + \frac{1}{2} t_2 \Phi \rho_{xy}C^2, \]

The MSE of \(\hat{Y}_{RDS}\) is given by

\[ MSE(\hat{Y}_{RDS}) \equiv \left( t_1 - 1 \right)^2 t_2 Y^2 + t_1 t_2 \Phi^2 A + t_1 t_2 \sum \phi X^2 B - t_1 t_2 \Phi^2 C. \]

Two auxiliary variables

(i) The traditional ratio-ratio estimator using two auxiliary variables is given by

\[ \hat{Y}_{RR} = \sum \phi \left( \frac{X}{X_{sys}} \right) \left( \frac{Z}{Z_{sys}} \right), \]

The bias and MSE of \(\hat{Y}_{RR}\) to first order of approximation are respectively given by

\[ B(\hat{Y}_{RR}) \equiv \Phi \left\{ \rho_{xy} C^2 + \rho_{xy} C^2 + C_y \sqrt{\rho_{xy} \rho_x} - C_y \sqrt{\rho_{xy} \rho_x} \right\}, \]

and

\[ MSE(\hat{Y}_{RR}) \equiv \Phi \left\{ \rho_{xy} C^2 + \rho_{xy} C^2 + \rho_{xy} C^2 + 2C_y \sqrt{\rho_{xy} \rho_x} - 2C_y \sqrt{\rho_{xy} \rho_x} \right\}. \]
\[ \hat{Y}_{RP} = \bar{y}^{(ys)} \left( \frac{X}{x^{(ys)}} \right)^{\frac{x^{(ys)}}{Z}} \]  

...(19)

(iii) The bias and MSE of \( \hat{Y}_{RP} \) to first order of approximation are given by

\[ B(\hat{Y}_{RP}) = \Phi \Phi \left\{ \rho_i^*C_i^2 - C_{xy}\sqrt{\rho_i^*\rho_i} - C_{xy}\sqrt{\rho_i^*\rho_i} + C_{xy}\sqrt{\rho_i^*\rho_i} \right\} \]  

...(20)

and

\[ MSE(\hat{Y}_{RP}) = \Phi \Phi \left\{ \rho_i^*C_i^2 + \frac{1}{4}\rho_i^*C_i^2 + \frac{1}{4}\rho_i^*C_i^2 + \frac{1}{2}C_{xy}\sqrt{\rho_i^*\rho_i} \right\} \]  

...(21)

(iv) The exponential ratio-ratio type estimator is given by

\[ \hat{Y}_{RRE} = \bar{y}^{(ys)} \exp \left( \frac{X - x^{(ys)}}{X + x^{(ys)}} \right) \exp \left( \frac{Z - z^{(ys)}}{Z + z^{(ys)}} \right) \]  

...(22)
\[
\hat{Y}_{QKH} = \frac{1}{n} \sum_{i=1}^{n} \left( X_i - \bar{X} \right) \exp \left( \alpha \frac{Z_i - \bar{Z}^{(QKH)}}{Z_i + (v_1 - 1)\bar{Z}^{(QKH)}} \right), \tag{30}
\]

where \( v_1 \) and \( v_2 \) are constants and \( \alpha \) is the scalar quantity, which takes values \((0, -1, +1)\).

The bias of \( \hat{Y}_{QKH} \) reported by Qureshi et al. (2018) is given by
\[
B(\hat{Y}_{QKH}) = \frac{1}{n} \sum_{i=1}^{n} \left[ v_i C_i^2 \left( 1 - H_{i-1} \right) \sqrt{\rho^2_{iy} + \frac{\alpha}{2v_i} C_i \left( 2(\alpha - 1) + \alpha - 2\sqrt{\rho^2_{iy} - v_i H_{i-1} \sqrt{\rho^2_{iy}}} \right)} \right],
\]

where \( H_{i-1} = \rho_{iy} C_i / C_j \), \( \rho^*_y = 1 + (n-1)\rho_y \), \( a = \alpha / \alpha^* \), and \( \alpha^* = \sqrt{\rho^2_{iy} - v_i H_{i-1} \sqrt{\rho^2_{iy}}} / (1 - \rho^2_{iy}) \).

Note: It is observed that the bias expression of estimator \( \hat{Y}_{QKH} \) by Qureshi et al. (2018) is not correct. Using equation (30), the correct bias expression is:
\[
B(\hat{Y}_{QKH}) \simeq \Phi \left[ \frac{1}{v_i} C_i \sqrt{\rho^2_{iy} + \frac{\alpha}{2v_i} C_i \left( 1 + \alpha - 2\sqrt{\rho^2_{iy} - v_i H_{i-1} \sqrt{\rho^2_{iy}}} \right)} \right] - \frac{1}{2} \left[ 1, (\alpha / v_i) \right.
\]

The minimum MSE of \( \hat{Y}_{QKH} \) at optimum values of \( v_i \)

and \((\alpha / v_i)\) i.e. \( \left( v_i \right)_{opt} = \frac{C_i \sqrt{\rho^2_{iy} - \rho^2_{iy} \rho_{iy} \rho_{iy} \rho_{iy}}}{C_i \sqrt{\rho^2_{iy} - \rho^2_{iy} \rho_{iy} \rho_{iy}} - 1} \) and \( \left( \alpha / v_i \right)_{opt} = \frac{C_i \sqrt{\rho^2_{iy} - \rho^2_{iy} \rho_{iy} \rho_{iy} \rho_{iy}}}{C_i \sqrt{\rho^2_{iy} - \rho^2_{iy} \rho_{iy} \rho_{iy}}} \) is given by
\[
MSE(\hat{Y}_{QKH})_{min} = \Phi \left[ \frac{\rho^2_{iy} C^2_i}{1 - R^2_{iy}} \right], \tag{32}
\]

(ix) Mishra et al. (2018) suggested the following class of estimators:
\[
\hat{Y}_{MSS} = \delta_0 \hat{Y}_0 + \delta_1 \hat{Y}_{RP} + \delta_2 \hat{Y}_{RPE}, \tag{33}
\]

where \( \sum_{i=0}^{2} \delta_i = 1 \) are constants; \( \hat{Y}_i (i = 0, RP, RPE) \) are defined earlier.

Mishra et al. (2018) derived the bias and minimum MSE of \( \hat{Y}_{MSS} \) up to first order of approximation as:
\[
B(\hat{Y}_{MSS}) \simeq \Phi \left[ \left( \delta_0 + \frac{\delta_0}{2} \right) \left( C_{sy} \sqrt{\rho^2_{iy} - C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}} \rho_{iy} \rho_{iy}} \right) \right.
\]

and
\[
MSE(\hat{Y}_{MSS})_{min} \simeq \Phi \left[ \rho^2_{iy} C^2_i + 3 \left( C_{sy} \sqrt{\rho^2_{iy} \rho^2_{iy} - C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}}} \right)^2 \right], \tag{35}
\]

The optimum value of \( \delta_0 \) is:
\[
\left( \delta_0 + \frac{\delta_0}{2} \right)_{opt} = \frac{\left( C_{sy} \sqrt{\rho^2_{iy} \rho^2_{iy} - C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}}} \right)}{\rho^2_{iy} C^2_i + \rho^2_{iy} C^2_i - 2C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}}} \tag{36}
\]

Expressions given in equations (34), (35) and (36) are not correct.

Solving equation (33), the correct expressions of bias and MSE of estimator \( \hat{Y}_{MSS} \) to first order of approximation are:
\[
B(\hat{Y}_{MSS}) \simeq \Phi \left[ \left( \delta_0 + \frac{\delta_0}{2} \right) \left( C_{sy} \sqrt{\rho^2_{iy} \rho^2_{iy} - C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}}} \right) \right.
\]

and
\[
MSE(\hat{Y}_{MSS})_{min} \simeq \Phi \left[ \rho^2_{iy} C^2_i \left( C_{sy} \sqrt{\rho^2_{iy} - C_{xy} \sqrt{\rho^2_{iy} \rho^2_{iy}} \rho_{iy} \rho_{iy}} \right)^2 \right], \tag{38}
\]

The correct optimum value of \( \delta_0 \) is:
\[
\left( \delta_0 + \frac{\delta_0}{2} \right)_{opt} = \left( \delta_0 + \frac{\delta_0}{2} \right)_{opt} \tag{39}
\]

**Proposed estimator**

Motivated by Khan (2016) and Riaz et al. (2017), the following difference-cum-exponential ratio type estimator is proposed for the population mean under systematic sampling:
\[ \hat{Y}_{n} = \left\{ A_{1}(\theta^{(1)}) + A_{2}(X - \bar{X}^{(1)}) \right\} + A_{3}(Z - \bar{Z}^{(1)}) \right\} \exp \left( \frac{X - \bar{X}^{(1)}}{X + \bar{X}^{(1)}} \right) \]

where \( A_{i} (i = 1, 2, 3) \) are constants whose values are to be determined.

The bias of \( \hat{Y}_{n} \) to first order of approximation is given by

\[ B(\hat{Y}_{n}) \equiv (A_{1} - 1) \bar{Y} + A_{2} \bar{Y} \Phi_{n} \left\{ 3 \rho_{1}^{2} C_{1}^{2} - 2 \rho_{1}^{4} \right\} \]

Substituting the optimum values of \( A_{i} (i = 1, 2, 3) \) in equation (42), we get the minimum MSE of \( \hat{Y}_{n} \) given by

\[ \text{MSE}(\hat{Y}_{n})_{\text{min}} \equiv \]

\[ \frac{1}{\alpha_{1}} \left[ \frac{1}{\Phi_{n}^{2}} \left( 1 - R_{y,x}^{2} \right) + \frac{1}{\Phi_{n}^{2}} \left( 1 - R_{y,x}^{2} \right) \alpha_{2} \right] > 0 \]

where \( \alpha_{2} = \frac{1}{4} \Phi_{n}^{2} C_{y}^{2} \left( 1 - R_{y,x}^{2} \right) \) and

\[ \alpha_{3} = \Phi_{n}^{2} \left( \rho_{2}^{2} - \rho_{y,x}^{2} \right)^{2} + \alpha_{2} + \alpha_{3} \]

Comparison of estimators

We compare the proposed estimator \( \hat{Y}_{n} \) with the following estimators

\( \hat{Y}_{g} \) (i = 0, R, RE, (D, KC), RDS, RR, RP, RRE, RPE, (DD, QKH), MSS(c), Pr) as:

(i) By equations (2) and (43), \( \text{MSE}(\hat{Y}_{n})_{\text{min}} < \text{Var}(\hat{Y}_{0}) \) if

\[ \frac{\text{MSE}(\hat{Y}_{n})}{\alpha_{1}} \left[ \Phi_{n}^{2} \left( 1 - R_{y,x}^{2} \right) + \Phi_{n}^{2} \left( 1 - R_{y,x}^{2} \right) \alpha_{2} \right] > 0 \]

where \( \alpha_{2} = \frac{1}{4} \Phi_{n}^{2} C_{y}^{2} \left( 1 - R_{y,x}^{2} \right) \) and

\[ \alpha_{3} = \Phi_{n}^{2} \left( \rho_{2}^{2} - \rho_{y,x}^{2} \right)^{2} + \alpha_{2} + \alpha_{3} \]

(ii) By equations (5) and (43), \( \text{MSE}(\hat{Y}_{n})_{\text{min}} < \text{MSE}(\hat{Y}_{g}) \) if

\[ \frac{\text{MSE}(\hat{Y}_{n})}{\alpha_{1}} \left[ \Phi_{n}^{2} \left( 1 - R_{y,x}^{2} \right) + \Phi_{n}^{2} \left( 1 - R_{y,x}^{2} \right) \alpha_{2} \right] > 0 \]

where \( \alpha_{2} = \frac{1}{4} \Phi_{n}^{2} C_{y}^{2} \left( 1 - R_{y,x}^{2} \right) \) and

\[ \alpha_{3} = \Phi_{n}^{2} \left( \rho_{2}^{2} - \rho_{y,x}^{2} \right)^{2} + \alpha_{2} + \alpha_{3} \]

From equation (42), the optimum values of \( A_{i} (i = 1, 2, 3) \) are

\[ A_{1(\text{opt})} = \frac{1 - \frac{1}{8} \Phi_{n} C_{y}^{2}}{1 + \frac{1}{4} \Phi_{n} C_{y}^{2} \left( 1 - R_{y,x}^{2} \right)} \]

\[ A_{2(\text{opt})} = \]

\[ \frac{1}{\phi_{n}} \left[ \frac{1}{2} C_{y} \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) - \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) \right] + A_{3} \]

\[ \frac{1}{\phi_{n}} \left[ \frac{1}{2} C_{y} \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) - \left( \rho_{2} C_{y} - \rho_{x}^{2} \right) \right] + A_{3} \]

and
(iii) By equations (8) and (43), \( \text{MSE}(\hat{Y}_{pr})_{\text{min}} < \text{MSE}(\hat{Y}_{rr}) \) if
\[
\frac{\hat{Y}_3}{\alpha_1} \left[ \Phi \left( \frac{1}{2} \sqrt{\rho_i^* C_z - \sqrt{\rho_y^* C_y, \rho_y^*}} \right) + \alpha_z + \alpha_1 + \alpha_6 \right] > 0,
\]
where \( \alpha_3 = \frac{\text{MSE}(\hat{Y}_{rr})}{\hat{Y}_3^2} \Phi \rho_y^* C_y \left( 1 - R_{y, x}^2 \right) \).

(iv) By equations (10), (12) and (43), \( \text{MSE}(\hat{Y}_{pr})_{\text{min}} < \text{MSE}(\hat{Y}_{KC})_{\text{min}} \), if
\[
\frac{\hat{Y}_3}{\alpha_1} \left[ \alpha_z + \alpha_3 + \alpha_6 \right] > 0,
\]
where \( \alpha_6 = \frac{\text{MSE}(\hat{Y}_{pr})}{\hat{Y}_3^2} \Phi \rho_y^* C_y \left( 1 - R_{y, x}^2 \right) \).

(v) By equations (15) and (43), \( \text{MSE}(\hat{Y}_{pr})_{\text{min}} < \text{MSE}(\hat{Y}_{KC})_{\text{min}} \) if
\[
\frac{\hat{Y}_3}{\alpha_1} \left[ 1 + \left( \Phi \rho_y^* C_y \frac{1}{8} \right) \right] + \frac{\left( \frac{1}{2} \Phi \rho_y^* C_y, \Phi \right)}{\alpha_1} \left( 1 - R_{y, x}^2 \right) - 1 \left[ 0,
\]
where \( \alpha_3 = \frac{\text{MSE}(\hat{Y}_{pr})}{\hat{Y}_3^2} \Phi \rho_y^* C_y \left( 1 - R_{y, x}^2 \right) \) and
\( \alpha_6 = \Phi \rho_y^* C_y \left( R_{y, x} - \rho_y^* - \rho_z^* \right) + 2 \Phi \rho_y^* C_y \sqrt{\rho_y^*} \).

(vi) By equations (18) and (43), \( \text{MSE}(\hat{Y}_{pr})_{\text{min}} < \text{MSE}(\hat{Y}_{RR}) \) if
\[
\frac{\hat{Y}_3}{\alpha_1} \left[ \Phi \left( \sqrt{\rho_y^* C_y - \sqrt{\rho_y^* C_y, \rho_y^*}} \right) + \Phi \left( \sqrt{\rho_y^* C_y - \sqrt{\rho_y^* C_y, \rho_y^*}} \right) \right] + \alpha_z + \alpha_1 + \alpha_6 \geq 0,
\]
where \( \alpha_3 = \frac{\text{MSE}(\hat{Y}_{pr})}{\hat{Y}_3^2} \Phi \rho_y^* C_y \left( 1 - R_{y, x}^2 \right) \) and
\( \alpha_6 = \Phi \rho_y^* C_y \left( R_{y, x} - \rho_y^* - \rho_z^* \right) + 2 \Phi \rho_y^* C_y \sqrt{\rho_y^*} \).

(vii) By equations (21) and (43), \( \text{MSE}(\hat{Y}_{pr})_{\text{min}} < \text{MSE}(\hat{Y}_{RR}) \) if
\[
\frac{\hat{Y}_3}{\alpha_1} \left[ \Phi \left( \sqrt{\rho_y^* C_y - \sqrt{\rho_y^* C_y, \rho_y^*}} \right) + \Phi \left( \sqrt{\rho_y^* C_y + \sqrt{\rho_y^* C_y, \rho_y^*}} \right) \right] + \alpha_z + \alpha_3 \left( 2 \Phi \rho_y^* C_y \sqrt{\rho_y^*} \right) \geq 0,
\]
where \( \alpha_3 = \frac{\text{MSE}(\hat{Y}_{pr})}{\hat{Y}_3^2} \Phi \rho_y^* C_y \left( 1 - R_{y, x}^2 \right) \) and
\( \alpha_6 = \Phi \rho_y^* C_y \left( R_{y, x} - \rho_y^* - \rho_z^* \right) - 2 \Phi \rho_y^* C_y \sqrt{\rho_y^*} \).

Numerical examples

We use the following 4 datasets from different sources.

**Data 1:** [Source: Mukhopadhyay (1998)]

Let \( y \) = output of factories, \( x \) = number of workers and \( z \) = fixed capital.
The data statistics are:

\[ \begin{align*}
N &= 40, & n &= 10, & \bar{Y} &= 5078.575, \\
\bar{X} &= 230.325, & \bar{Z} &= 945.425, & C_y &= 0.3295, \\
C_x &= 0.8465, & C_z &= 0.6755, & \rho_{xy} &= 0.80056, \\
\rho_{xz} &= 0.8349, & \rho_{yx} &= 0.9816, & \rho_{yz} &= 0.23, \\
\rho_{zx} &= 0.086, & \rho_{zy} &= 0.16. \\
\end{align*} \]

**Data 2:** [Source: Gujrati (2012)]

Let \( y \) = average miles per gallon (MPG), \( x \) = top speed, miles per hour (TS) and \( z \) = engine horsepower (HP).

The passenger car mileage data statistics are:

\[ \begin{align*}
N &= 81, & n &= 21, & \bar{Y} &= 33.84, \\
\bar{X} &= 112.4568, & \bar{Z} &= 177.4691, & C_y &= 0.29719, \\
C_x &= 0.12555, & C_z &= 0.48620, & \rho_{xy} &= -0.6907, \\
\rho_{xz} &= -0.7944, & \rho_{yx} &= 0.9068, & \rho_{yz} &= 0.56, \\
\rho_{zx} &= 0.35, & \rho_{zy} &= 0.52. \\
\end{align*} \]

**Data 3:** [Source: Murthy (1967)]

Let \( y \) = area under wheat (in acres) during 1974, \( x \) = area under wheat (in acres) during 1971 and \( z \) = area under wheat (in acres) during 1973.

The data statistics are:

\[ \begin{align*}
N &= 34, & n &= 9, & \bar{Y} &= 199.4412, \\
\bar{X} &= 856.411, & \bar{Z} &= 208.8824, & C_y &= 0.7531, \\
C_x &= 0.8560, & C_z &= 0.7205, & \rho_{xy} &= 0.98, \\
\rho_{xz} &= 0.4491, & \rho_{yx} &= 0.11, & \rho_{yz} &= 0.091, \\
\rho_{zx} &= 0.043. \\
\end{align*} \]

**Data 4:** [Source: Tailor et al. (2013) and Mishra et al. (2018)]

The hypothetical data statistics are:

\[ \begin{align*}
N &= 15, & n &= 3, & \bar{Y} &= 80, \\
\bar{X} &= 44.47, & \bar{Z} &= 48.40, & C_y &= 0.56, \\
\end{align*} \]

<table>
<thead>
<tr>
<th>Table 1: Percent relative efficiency (PRE) of different estimators with respect to ( \hat{Y}_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimator</td>
</tr>
<tr>
<td>( \hat{Y}_0 )</td>
</tr>
<tr>
<td><strong>One auxiliary variable</strong></td>
</tr>
<tr>
<td>( \hat{Y}_R )</td>
</tr>
<tr>
<td>( \hat{Y}_{RE} )</td>
</tr>
<tr>
<td>( \hat{Y}<em>{DR}, \hat{Y}</em>{KC} )</td>
</tr>
<tr>
<td>( \hat{Y}_{RDS} )</td>
</tr>
<tr>
<td><strong>Two auxiliary variables</strong></td>
</tr>
<tr>
<td>( \hat{Y}_{RR} )</td>
</tr>
<tr>
<td>( \hat{Y}_{RP} )</td>
</tr>
<tr>
<td>( \hat{Y}_{RE} )</td>
</tr>
<tr>
<td>( \hat{Y}_{RPE} )</td>
</tr>
<tr>
<td>( \hat{Y}<em>{DPE}, \hat{Y}</em>{QKH} )</td>
</tr>
<tr>
<td>( \hat{Y}_{MSS(c)} )</td>
</tr>
<tr>
<td>( \hat{Y}_{Pr} )</td>
</tr>
</tbody>
</table>
The following expression is used to obtain the percent relative efficiency (PRE):

\[
PRE = \frac{\text{Var}(\hat{\sigma}_i)}{\text{MSE}(\hat{\sigma}_i) \text{ or } \text{MSE}(\hat{\sigma}_i)_{\text{min}}} \times 100,
\]

\(i = 0, R, RE, (D, KC), RDS, RR, RP, RRE, RPE, (DD, QKH), \text{ MSS(c)} \) and \( \text{Pr} \).

The \( PRE \) values for different estimators based on the above datasets are given in Table 1.

**RESULTS AND DISCUSSION**

From Table 1, it is observed that the performance of the proposed estimator (\( \hat{\sigma}_p \)) is the best among all the estimators considered here. The estimators \( \hat{\sigma}_e, \hat{\sigma}_a, \hat{\sigma}_b, \hat{\sigma}_{RP}, \hat{\sigma}_{RRE} \) show poor performances for some datasets, i.e. (Data 1, Data 2), Data 2, (Data 1, Data 2, Data 4), (Data 1, Data 3), (Data 1, Data 2, Data 4) and Data 3, respectively. The performance of difference estimator (\( \hat{\sigma}_d \)) and Kocyigit and Cingi (2017) estimator (\( \hat{\sigma}_{KCY} \)) are equal when using single auxiliary variable. The improvement in Riaz et al. (2017) estimator (\( \hat{\sigma}_{RYS} \)) is more than (\( \hat{\sigma}_d \)) and (\( \hat{\sigma}_{KCY} \)) estimators. Also difference estimators (\( \hat{\sigma}_{RYS}, \hat{\sigma}_{RYSK} \)) are equally efficient but (\( \hat{\sigma}_{KCY} \)) is preferable because of unbiasedness for using two auxiliary variables.

**CONCLUSION**

A new improved difference-cum-exponential ratio type estimator is proposed for estimating the finite population mean under systematic sampling using single and two auxiliary variables. Some estimators, when using two auxiliary variables do not perform better than the estimators using single auxiliary variable. The proposed estimator (\( \hat{\sigma}_p \)) is more efficient than the usual sample mean estimator (\( \hat{\sigma}_d \)), traditional ratio estimators (\( \hat{\sigma}_e, \hat{\sigma}_a \)), exponential-ratio estimators (\( \hat{\sigma}_e, \hat{\sigma}_a \)), ratio-product estimators (\( \hat{\sigma}_{RPE}, \hat{\sigma}_{RPEK} \)), difference estimators (\( \hat{\sigma}_d, \hat{\sigma}_{KCY} \)) and the other estimators considered here.

**REFERENCES**


DOI: https://doi.org/10.1515/eqc-2013-0005


DOI: https://doi.org/10.22237/jmasm/1320120180


DOI: https://doi.org/10.26438/ijsrmss/v5i3.7378